N63-85845

NASA CR 51962

ENGINEERING RESEARCH LABORATORIES

COLLEGE OF ENGINEERING
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA

REPORT ON TELEMETRY, CODING AND DATA PROCESSING SYSTEMS FOR THE ASTRONOMICAL ORBITING OBSERVATORIES PROJECT

Prepared for

THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

By

E. C. BELLEE, J. W. BLANCHARD, G. R. PETERSON and E. L. MORRISON

(NASA CR-51968)

Of

0727602

**THE APPLIED RESEARCH LABORATORY
THE UNIVERSITY OF ARIZONA
TUCSON, ARIZONA**

TABLE OF CONTENTS

TABLE OF ILLUSTRATIONS

II. SATELLITE DESCRIPTION

In order to establish a frame of reference for the following discussions of the telemetering system, it seems logical to briefly review the overall organization of the various satellites. Some of the readers of this report may be familiar only with certain parts of the program, and even those persons familiar with the overall program may find a brief review helpful in following the ensuing discussions.

As the program is now planned, all the experiments will ride the same stabilized platform. Although this platform has not as yet reached a final design stage, its general organization can be specified. The platform will include facilities for stabilization, guidance, power supply, and a finder system.

The stabilization and guidance system will probably involve stabilizing jets in three axes, reaction wheels in three axes, and probably rate gyros in three axes. The control for the guidance system will probably be a television camera viewing a small field, and acting in a closed loop system to hold a particular star in some very small area of the field. In the case of the solar experiments it might be necessary to use some form of inertial guidance, in which case position gyros in three axes will be required.

The power supply system will probably consist of solar cells charging some sort of storage batteries, The finder system will consist of a television camera viewing the field of a small telescope, and transmitting the image back to earth, where it will be interpreted by an astronomer or a computer. One additional requirement will be for a number of temperature sensors,to report on thermal conditions on the satellite.

As to the astronomical elements of the various experiments, only the Princeton and Smithsonian experiments will be considered. These two are the most complex of the various experiments, and any telemetering system adequate for them should be more than adequate for the other experiments.

The basic element in the Princeton unit will be a 24" telescope with a slit spectroscope with 0.1A resolution for studies of stellar spectra in the ultra-violet region. The transducers for the spectroscope will be five photo-multiplier tubes with pulse outputs. This satellite will be on a polar orbit, placing it in transmission range for 15 minutes, once every twelve hours. During the twelve hours that the unit is out of sight it will automatically scan a region of the spectrum in 0.1A steps, with an exposure time of 20 to 160 seconds in each interval. The pulses from the tube will be counted, and the total count for each interval, together with the identification of the interval will be stored. The maximum number of bits to be stored in a 12 hour period will be 180,000, all of which must be transmitted back to earth in the 15 minute period that the satellite is in sight.

The Smithsonian unit will consist of a cluster of five 8" telescopes. Three of these will be viewed by slow-scan TV cameras, to make a sky survey in three regions of the ultraviolet. The fourth scope will have a slitless spectroscope and will also be viewed by a TV camera. The fifth scope will be equipped with a photomultiplier tube for photometric intensity measurements. This satellite will be on an orbit inclined at 35° to the equator, and will be tracked by two stations in the United States. With this arrangement the satellite will be in transmission range for approximately 30 minutes out of every 100 minutes. There will be no storage requirements for this unit, as all measurements will be made while the satellite is in transmission range, the unit being inactive during the remainder of the orbit.

The function and modes of operation of the telemetry and communication systems for the satellite can best be illustrated by a summary of the events taking place from the time of the launch to the time that the satellite is actually gathering data. This summary is presented in outline form below and is the interpretation of the system operation derived from the published minutes of the two meetings on the O. A. O. project.

LAUNCH TO ORBIT

1. Telemetry System Off

   It is assumed that during the launch all telemetering will be handled by the launch vehicle. However, the satellite telemetry system must be switched to a standby operation ready to receive commands prior to separation from the launching vehicle.

2. Minitrack System On

   The minitrack transmitter may be off during the initial launch but must be activated immediately upon separation of the satellite and lauach vehicle to allow ground tracking of the satellite to determine when a stable orbit has been attained.

ORBIT

A. Initial Stabilization

1. Telemetry System on Command and Guidance

   The telemetry system would be enabled from the ground and the course stabilizing program started. Telemetry data transmitted would only be that pertaining to guidance, control, and power system operation.

2. Data Television System Off

   It is assumed that during this phase of the stabilization, the angular rate of the platform will be too great for the TV to be of any use. Information will come from rate gyros, etc., which will be handled by the regular telemetry system.

B. Final Stabilization and Orientation

1. Telemetry System on Command and Guidance

Assume the stabilizing program has progressed to the final stabilization and the sensors of interest here are different than those used in the initial stabilization phase. If sufficient channels are available, this telemetry system mode can be the same as the previous one and no change-over will be necessary.

2. Fine Television System Off

3. Finder Television System On

The wide angle finder television system will be used for the orientation of the platform and probably for the final stabilization.

C. Data Gathering Phase

1. Fine television System On

The fine television system would also be used in the final stages of the platform orientation to insure that the desired image is centered on the TV viewing field.

2. Astronomical Data Program On

This program would enable the various transducers in the experiment and start the data collection.

3. Telemetry System on Astronomical Data and Command

Data telemetered in this phase should be astronomical data and command verifications concerning the astronomical equipment and its functions.

III. RADIO PROPAGATION CHARACTERISTICS AND FREQUENCY ALLOCATION CONSIDERATIONS

The purpose of this section of the report is to discuss briefly factors that must be considered in the choice of operating frequencies for the various communication systems in the satellite.

This discussion is based on the following general assumptions:

1. Earth-satellite communications are to be designed for communicating only when the satellite is within line-of-sight view of the ground stations.

2. The conservation of power on the satellite while important, is of secondary importance to considerations of reliability. Progress in power supply design and launch capabilities has been so great in recent years, that it is no longer necessary to worry about every milliwatt, as it was on the first satellites.

3. Parameters chosen should represent state-of-the-art and not projected values.

4. Communication signals to and from the satellite must pass through the earth's atmosphere. Frequency selective "windows" appear in this propagation medium. The lower window between approximately 10mc and 10kmc will be chosen because state-of-the-art equipment exists for communication in this frequency spectrum. This "window" is not sharply defined on either end, being limited on the lower end by ionospheric absorption and refraction and on the upper end by tropospheric (rainfall and gaseous) absorption.

5. Within the lower window "free space" propagation losses are given by the relationship

$$\frac{P_t}{P_r} = G_t \, G_r \left(\frac{\lambda}{4 d}\right)^2$$

P_t = power delivered to transmitting antenna

P_r = receiving antenna power gain

G_t = transmitting antenna power gain

λ = wavelength

d = distance between transmitting and receiver antennas

Discussion

The nature of the experiments being performed from the satellites require that communications be maintained under all possible orientations of the satellite. Although it is true in general that the satellite, once stabilized, will always point away from the earth, one of the most important phases of the program is the initial stabilization and acquisition phase, during which data must be transmitted to the ground and command signals sent to the satellite. Since the vehicle may be tumbling initially or at least may be randomly oriented, it is essential that a non-directional antenna be used for at least certain portions of the system.

The basic system of antennas to be used will be a factor in determining the optimum R. F. frequency. Figure I illustrates the frequency characteristics of three combinations of ground and satellite antennas that could be considered; isotropic-directive, isotropic-isotropic, and directive-directive. The factors discussed above dictate the choice of an isotropic antenna on the satellite, while a directive antenna is desirable on the ground, in order to increase the system gain. Thus the use of an isotropic-directive system, Curve II in the figure is indicated.

Having made the basic choice of antenna system the effects of noise on this system may be considered. Here several sources of noise must be considered. The basic types can be categorized as follows:

1. Terrestrial Noise

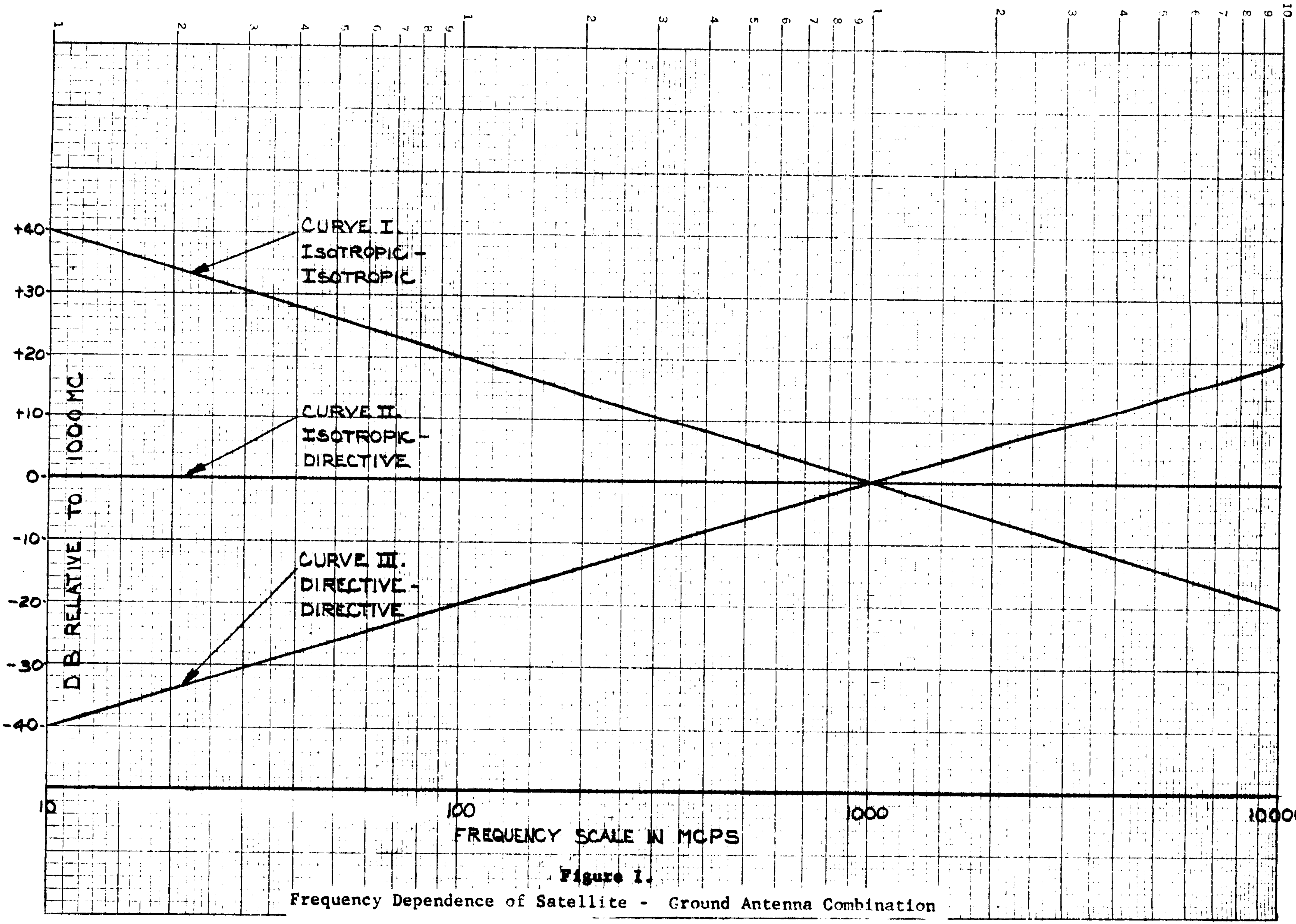


Figure I.

Frequency Dependence of Satellite - Ground Antenna Combination

A - Man made interference

B - Noise generated in the receiver equipment

2. Extra Terrestrial Noise

A - Solar noise

B - Galactic noise

Within each of the major categories there are many different sources that might be considered. However, many of these are insignificant for the present system and only the major sources of noise will be considered here.

If it is assumed that 10 mc is the lower limit of possible frequencies, then man-made sources of noise are not a major problem. Only the interference from coherent signal sources, such as harmonics from other transmitters, should be considered. This interference cannot be predicted, but must remain as a factor to be checked for any given location of the ground station.

The effective noise power of the receiver input terminals can be calculated by reference to the basic definitions of noise figure.

$$\text{Noise Figure} = \frac{(S_i/N_i)}{(S_o/N_o)} = \frac{\text{Signal-to-noise ratio at the input}}{\text{Signal-to-noise ratio at the output}}$$

or

$$\text{Noise Figure} = F = \frac{N_o}{k\ T_o\ BG} = \frac{\text{Actual Noise at the output}}{\text{Minimum Noise at the output}}$$

where

k = Boltzmann's constants = 1.374×10^{-23} joules per ^{o}K

T_o = $290^{o}K$

caused by some equivalent noise power (P_{ni}) at the input multiplied by the gain of the receiver (G). Then

$$F = \frac{P_{ni}G}{KT_oBG} = \frac{P_{ni}}{KT_oB}$$

Note here that P_{ni} consists of two components: (1) Johnson noise in the input termination (KT_oB); (2) the noise contributed by other factors in the system, (P_n). Thus

$$F = \frac{KT_oB + P_n}{KT_oB} = 1 + \frac{P_n}{KT_oB}$$

or

$$P_n = (F-1)\ KT_oB$$

The noise in the entire system can be considered to be due to an effective temperature such that

$$P_n = KT_eB = (F - 1)\ KT_oB$$

(1) $$T_e = (F-1)\ T_o$$

Here T_e is the "effective noise temperature" of the system, T_o is the reference temperature (290° K) and F is the overall noise figure of the combined system.

Noise figures for cascaded stages combine as

(2) $$F_{1-3} = F_1 + \frac{F_2-1}{G_1} + \frac{F_3-1}{G_1G_2}$$

where: F_{1-3} = overall noise figure of the 3 stage system

F_1 = overall noise figure of the 1st stage

F_2 = overall noise figure of the 2nd stage

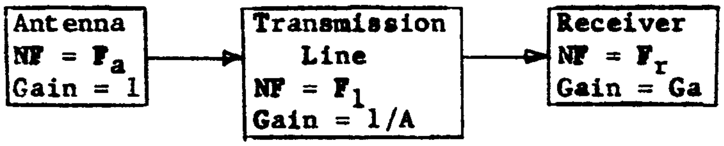
F_3 = overall noise figure of the 3rd stage

G_1 = overall power gain of the 1st stage

G_2 = overall power gain of the 2nd stage

G_3 = overall power gain of the 3rd stage

The system under consideration consists of the block diagram illustrated below:



The antenna contributes noise from the fact that its radiation resistance effectively assumes thermal equilibrium with whatever its aperature is pointed toward. To be precise this must include the effects of side lobes, back lobes, primary feed spillover, etc. However, these factors do not become significant except when masers are being considered. For this system the noise figure of the antenna may be written as

$$Fa = \frac{T_o + T_a}{T_o} = 1 + \frac{T_a}{T_o}$$

Ta = effective temperature of the source viewed by the antenna.

A lossy network has an effective noise figure because it serves to reduce the signal while leaving the noise level the same, hence the signal-to noise ratio at the output is reduced.

The noise figure is given by:

$$F_l = \frac{1}{\text{Gain}} = \text{Attenuation} = A$$

. The overall noise figure can now be determined by substitution in equation (2)

$$F_{1-3} = F_a + \frac{F_l - 1}{1} + \frac{(F_r - 1)}{1/A}$$

$$= \left(1 + \frac{T_a}{T_o}\right) + \frac{(A - 1)}{1} + (F_r - 1)\,A$$

(3) $$F_{1-3} = \frac{T_a}{T_o} + F_rA$$

The effective noise temperature of the system is given by equation (1) as

$T_e = (F_{1-3} - 1)\,T_o$ and substituting for F_{1-3}

(4) $$T_e = \left(\frac{T_a}{T_o} + F_rA - 1\right) T_o = T_a + (F_rA - 1)\,T_o$$

Equation (4) illustrates how the receiver noise figure, transmission line loss between the antenna and the receiver, and the noise contribution due to the antenna pointing at various black body radiators combine to cause an effective noise temperature of the receiving system.

Based on the above equation it is possible to plot the contributions of galactic noise and receiver characteristics in such a way as to predict at what frequency the minimum overall noise could be expected with present day equipment. To accomplish this result it is necessary to make assumptions concerning the available noise figure in state-of-the-art receivers and transmission line loss to be expected in a system configuration and how these

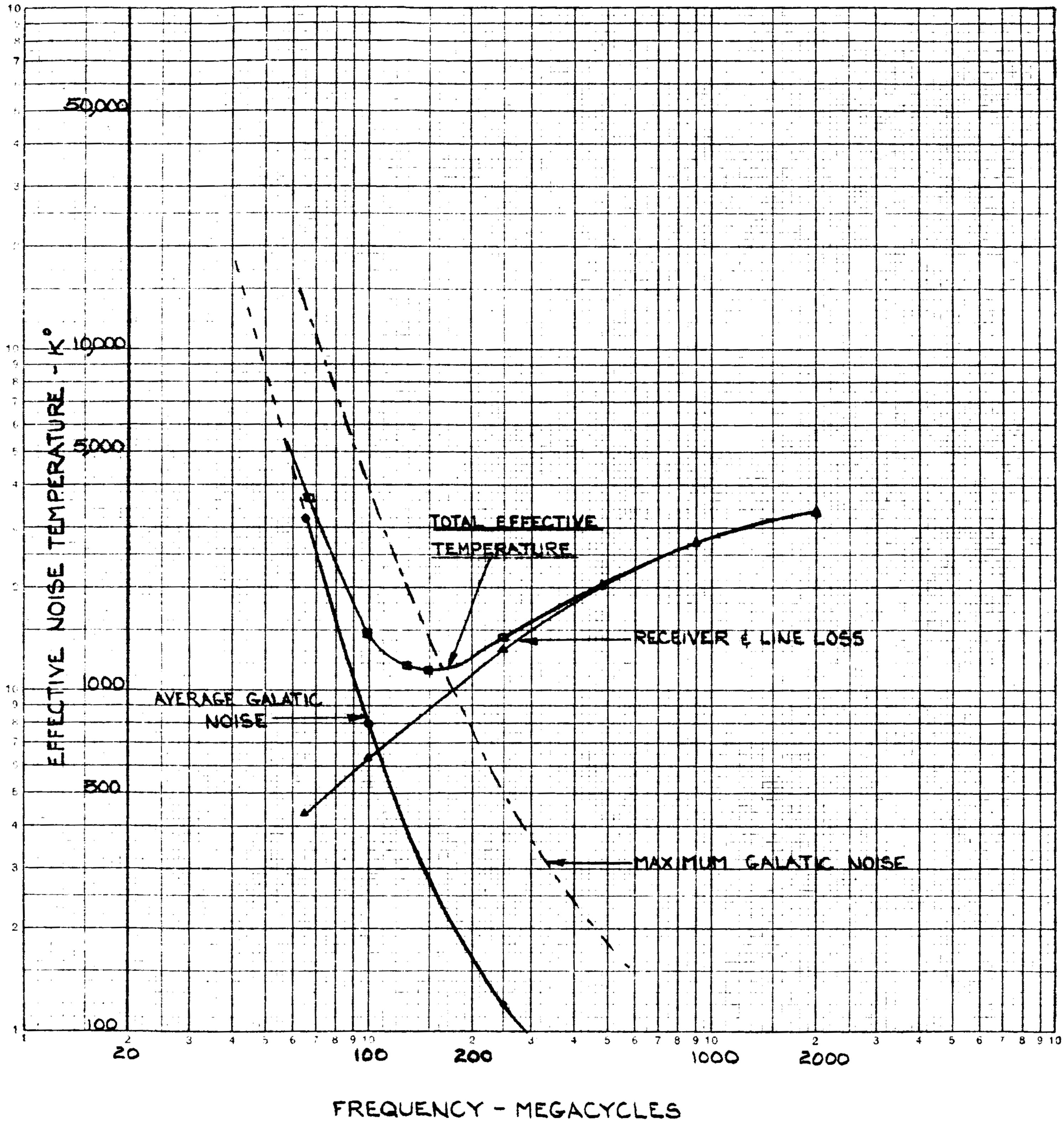


Figure II.
Effective Noise Temperature vs. Frequency

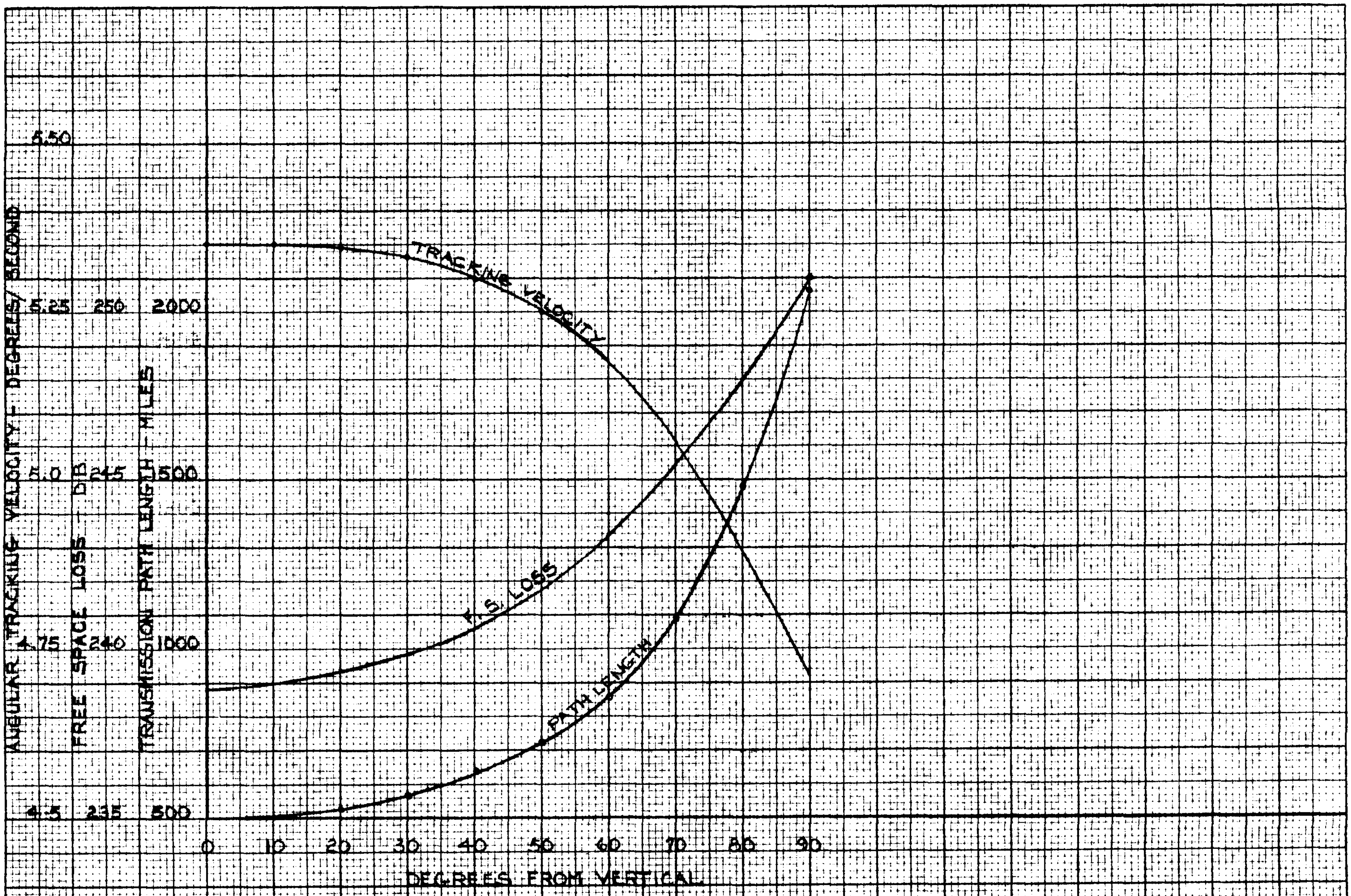
not sharp, however, and only a 3 DB increase in system noise should be expected up to approximately 2000 mc.

At this point it would seem that the operating frequency chosen should be near 150 mc and certainly this frequency affords many advantages. Nevertheless other systems aspects enter into the overall picture.

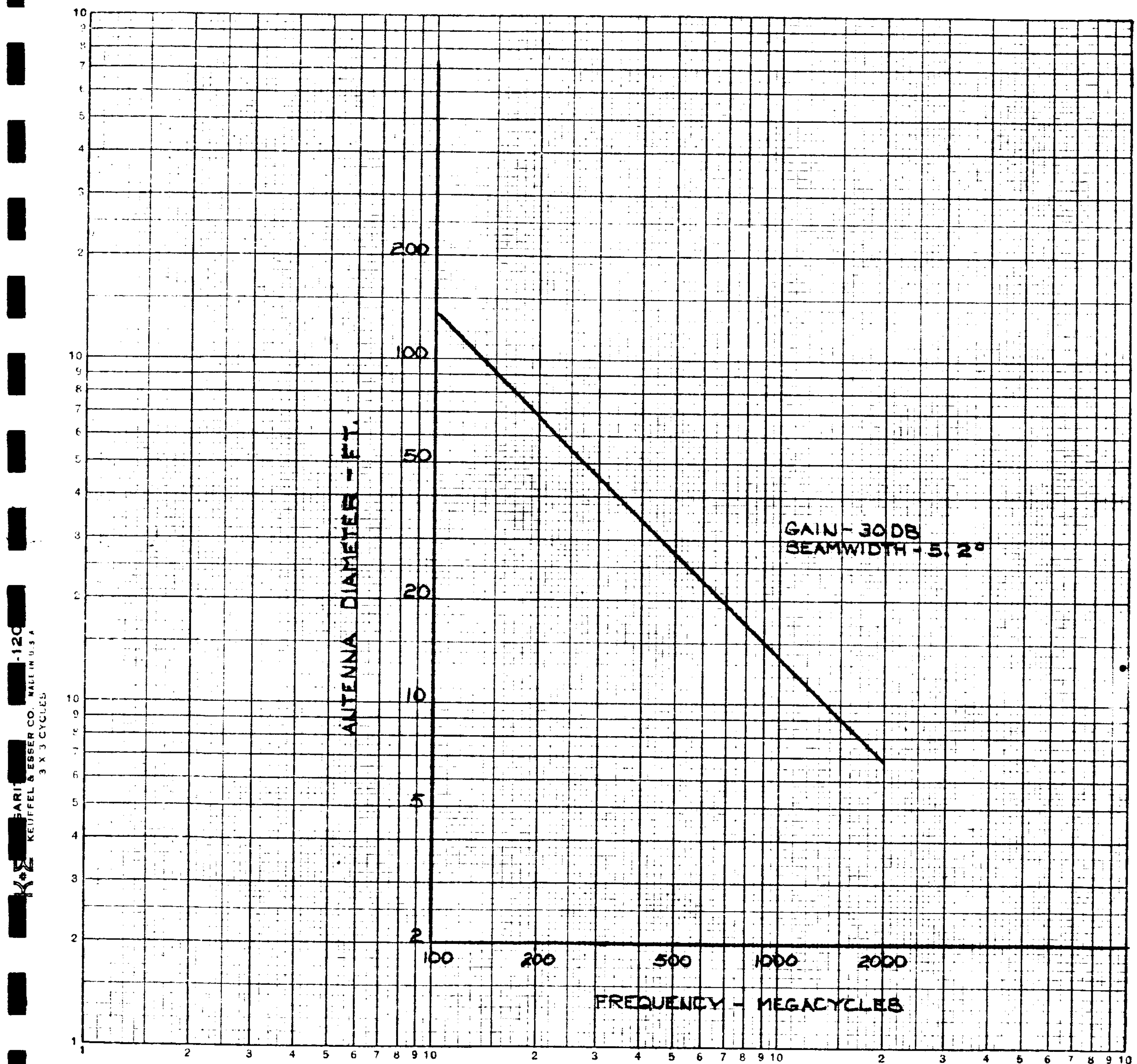
During operation with the satellite the antenna must be pointed such as to acquire the transmitted signal from the satellite as near the radio horizon as possible. After acquisition it must continue to track the satellite during the pass overhead. The satellite orbit will be known to the extent that with a reasonable bandwidth antenna it will be possible to point the antenna at the horizon where the satellite should appear and let automatic tracking take over as soon as the receiver signal strength reaches a certain minimum level. However, in the event of antenna tracking system failure it should still be possible for an operatorto manually track the satellite during the "pass". This is only possible with antenna beamwidths that are not too narrow.

FigureIII is a plot of path distance to the satellite and angular velocity of the satellite for a typical 500 mile circular orbit. It is to be noted that the maximum angular tracking velocity is shown as approximately 1/2 degree per sec. This implies that with an antenna beamwidth of 5^{o}, it would take only 10 seconds for the satellite to pass thru the beam of the antenna when it is vertically directed. Since the choice of antenna size is a compromise between beamwidth and gain, a reasonable middle-of-the-road figure appears to be the 5.25^{o} beamwidth, with an attendant 30 DB gain figure.

Having made this choice of antenna parameters, size as a function of frequency to achieve this beamwidth and gain (assuming a parabolic antenna) is plotted in Figure IV.



Orbital Characteristics - 500 Mile Circular Orbit
Figure III.



Frequency/Size Dependance-Parabolic Antenna

Figure IV.

The advantage of using the higher frequencies, 1000 mc and above, becomes very attractive in view of relative antenna size required.

If consideration is also given to future developments in communication, with the though in mind that equipment developed on this program will find further use in future space programs, then the future of the low-noise parametric amplifiers should be considered. It appears that within the very near future these amplifiers will be readily available, at least for frequencies above 1000 mc. When this is true the available noise figures used in plotting Figure II will no longer be valid. Instead the curve will show a drastic dip in this frequency range down to approximately 100° K and the choice of operating frequency will then obviously be wherever these amplifiers can be used. With this thought in mind, the most economical choice of operating frequency for the present equipment would appear to be near the high frequency telemetryband where advantage of the low noise amplifiers can be realized in the future without complete revision of the equipment. It is even possible that it may be economical to use this low noise amplifier by the time the equipment for this program is finally designed and built.

The actual choice of RF frequency for this satellite program must be governed by the economics of time and money. If time and money are available then choosing the RF frequency in the IRIG telemetry band of 2200-2300MC will offer many advantages to this and future systems of a similar nature. The disadvantage of this choice lies in the fact that the development of equipment for operation in this frequency band has lagged that in the lower telemetry frequency bands. The result is that considerable development money, and consequently time, would be necessary to realize working components for the satellite and ground systems.

The next most logical operating frequency would be approximately 275 MC. This frequency lies just above the IRIG Telemetry band of 216-260 MC. The 216-260 MC spectrum should not be used for this program because of the missle test range activity that is already overcrowding this band. Much equipment has been developed for operation in this part of the spectrum; engineering experience on equipment of this type is plentiful. A good share of the equipment necessary could be off-the-shelf equipment, since the range of operation could probably be made to extend to this frequency with little or no modification. The net result of choosing this frequency for operation would be to realize an operating system, compatible with present telemetry systems, at a reasonable cost and at an earlier date than might be possible with an operating frequency of 2200-2300MC.

The third choice of operating frequency would be near 150 MC*, at which frequency the minimum overall receiver noise is realized and hence minimum satellite power is required. The amount of power, and hence satellite weight, saved over that required for the other frequencies will not be great. However, the cost per pound of the satellite will be, and if the greatest overall economy is important, then it may be necessary to operate in this portion of the spectrum.

There are two other propagation factors that should be considered -- polarization and multipath effects. Multipath fading effects can be expected to be most deleterious when the satellite is near the horizon. At this time, grazing angles become favorable for efficient reflection to take place. This may result in large fluctuations in the signal strength as the phase of the

* Information received just prior to the final draft of the report indicates that this is the operating frequency chosen by NASA for the initial phase of the program.

reflected signal alternately reinforces and cancels the direct signal. A typical plot of this signal strength variation is shown in Figure V. The presence of multipath is characterized by a series of sharp nulls occurring in the signal strength and appearing as short bursts of noise in the receiver output as the signal level drops below system threshold.

Space diversity may be used to counteract these fluctuations if the attendant cable loss can be tolerated. However, if adequate antenna height is used, this fading will be no worse than the maximum of that for Rayleigh fading, for which allowance must be made in the system design in any case.

The satellite orientation must be considered as random, hence the polarization of the transmitted wave must be considered random. This means that the antenna system must be capable of receiving left circular, right circular, horizontally or vertically polarized waves. To realize efficient reception under these conditions of random polarization, polarization diversity should be used. Antennas with feed systems such as that used on the TLM-18 adapt very easily to a polarization diversity system. Under no operating conditons would the signal-to-noise ratio that is realized be worse than a system without diversity, and under equal signal conditions in the two channels 3 DB signal-to-noise ratio improvement can be obtained. In addition, polarization diversity can be expected to counteract, to a certain extent, the sharp nulls in signal strength shown in Figure V , since the location of the nulls is a function of the type of polarization being received. However, the extent of the improvement is impossible to predict with any accuracy.

Most of the fading that occurs on paths not subject to reflection is the result of interference between two or more rays traveling different paths in the atmosphere. This type of multipath effect is relatively independent

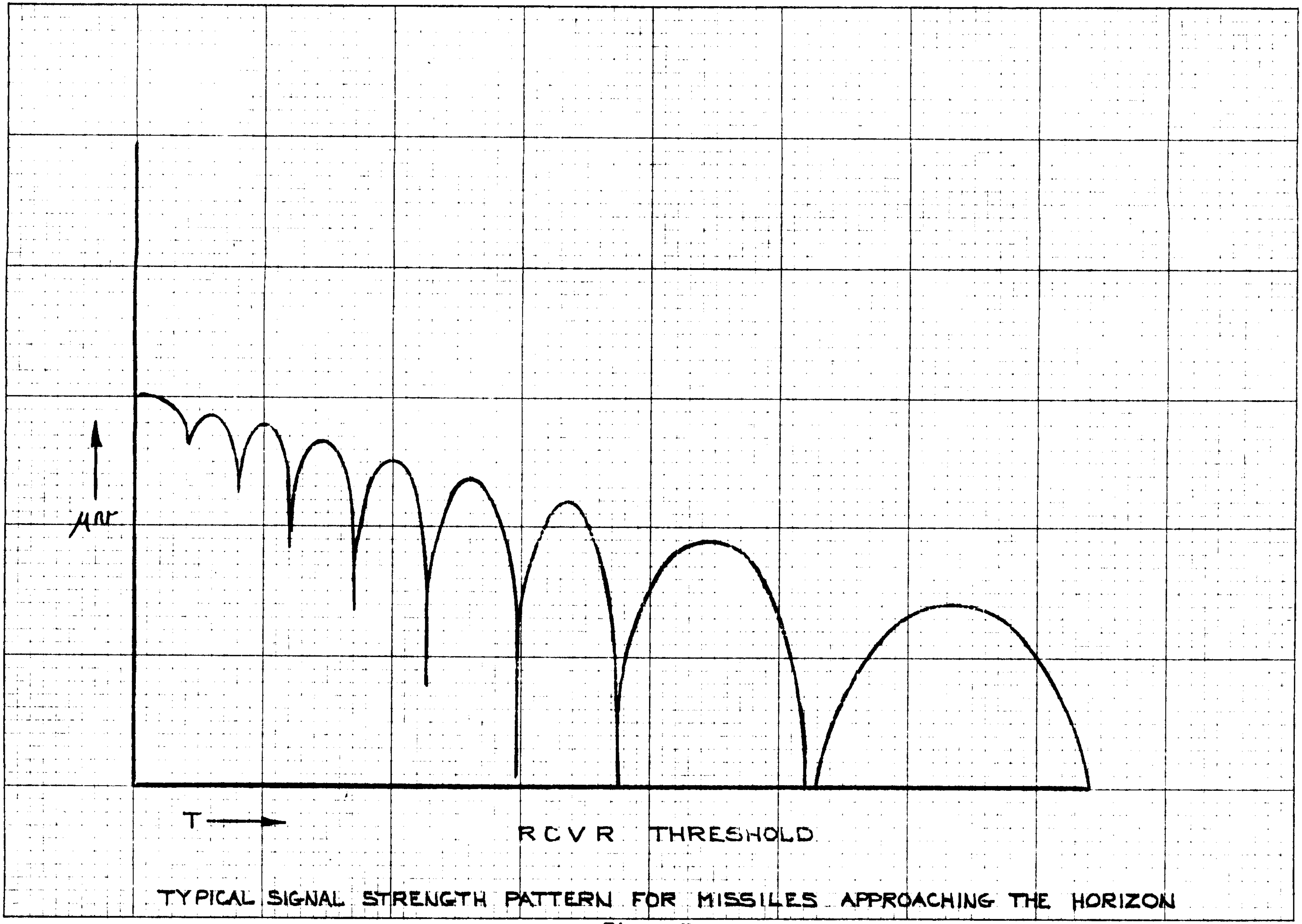

μv
T
RCVR THRESHOLD
TYPICAL SIGNAL STRENGTH PATTERN FOR MISSILES APPROACHING THE HORIZON

Figure V.

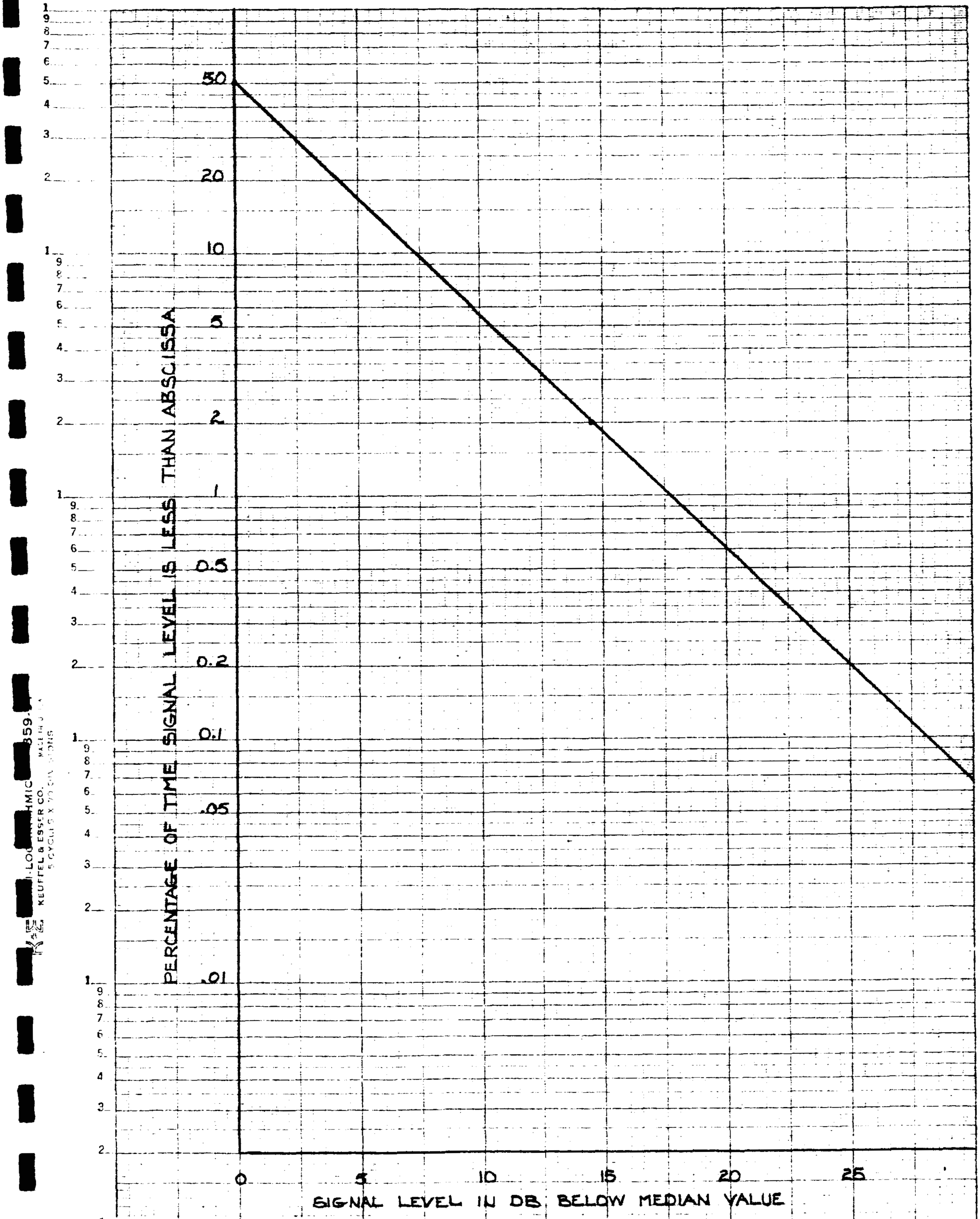
of the height of the path above the earth, and its extreme condition approaches a Rayleigh distribution, and hence is frequently referred to simply as "Rayleigh fading". In the Rayleigh distribution the probability that the instantaneous value of the field strength is greater than the value R is exp-(R/R_o), where R_o is the RMS value. This distribution is plotted in Figure VI.

In order to make allowance for multipath fading on a transmission link it is necessary to increase the transmitter output power by a factor that will provide the degree of transmission reliability desired. For instance, in the transmission circuit being considered, a 99 % reliability figure would be a reasonable choice. In this case, 1 % of the time the signal will be below the desired value. From the graph of the Rayleigh distribution this would require that an 18 DB fading margin be added to the RF power calculated from other considerations.

Conclusions

The recommended frequency of satellite operation is in the IRIG telemetry band of 2200-2300MC. In this spectrum high gain tracking antennas of small size may be used along with maser amplifiers to realize as much as a 10 to 1 reduction in satellite RF power over that required at other frequencies. Whether these advantages can be realized will be governed by the OAO Schedule and budget. Conceivably, the initial satellites may operate at the lower frequencies in order to meet their schedules, while the later and more complex satellites may be integrated into a system operating in the 2200-2300 MC range.

Polarization diversity should be used for the receiving antenna system if the attendant cost can be tolerated. While it does not materially decrease the satellite requirements it does provide additional communication reliability;



Rayleigh Distribution (Theoretical Maximum for Multipath Fading)
Figure VI.

a feature that is always desirable.

The fading margin allowance for this communication system is rather arbitrary. A good compromise, and the value recommended, is 18 DB. This value will provide a signal strength at the radio horizon which is above the minimum value 99% of the time. As the satellite continues its pass, climbing above the horizon, communication ability will improve beyond this value.

IV. MODULATION AND MULTIPLEXING

While it is beyond the scope of this report to analyze in detail the advantages and disadvantages of all the schemes of modulation and multiplexing that are available, a choice of a basic system can be made based on the work already done in this area. In order to do this it is necessary to first outline the requirements and objectives of the system.

Requirements and Objectives of the Telemetry System

1. The power requirements of the satellite will be to a large extent dependent upon the power requirements of the telemetry system used. Therefore, power efficiency and hence communication efficiency are important.

2. The satellite must have great longivity if the missions of the experimenters are to be realized. This means that the system must operate for an extended period of time without adjustments even under repeated turn on and turn off (standby) control of power to the unit.

3. To be abasic telemetry system for the O.A.O. satellite program, the system must have a great deal of flexibility. That is, it must be readily adaptable to various degrees of multiplexing, including both super and sub-communtation without major system design changes. This is necessary if the system is to adapt readily to the requirements of the different experiments.

4. Economically it is not advisable to undertake a completely new development in a program of this type unless it is absolutely necessary for successful completion of the project. In view of this, the basic telemetry system should be one that has been tested and proven wherever possible.

With these points in mind the various schemes of modulation and multiplexing may now be compared. A thorough analysis and comparison of these methods have been completed by Nichols and Rauch[61] and is tabulated in Table I., the following page.

Table I. Comparison of Telemetry Systems

Type	Threshold Carrier to Noise Ratio in Terms of AM Standard	RF Power (Relative to PPM/AM	Information Efficiency*	RF Bandwidth in KC for 1000 CPS Signal
1. PPM/AM	200	1.0	0.17	76
2. PCM/FM	260	1.7	0.24	18
3. PCM/PM	280	2.0	0.21	20
4. PCM/AM	370	3.4	0.21	18
5. PAM/FM	580	8.3	0.050	85
6. AM/FM	610	9.3	0.045	93
7. PDM/FM	610	9.3	0.045	92
8. PDM/PM	660	11.0	0.036	110
9. FM/FM	740	14.0	0.030	140
10. AM/PM	770	15.0	0.028	150
11. PAM/PM	780	15.0	0.028	150
12. PDM/AM	790	16.0	0.035	94
13. FM/AM	830	17.0	0.055	50
14. FM/PM	860	18.0	0.023	185
15. PAM/AM	3150	250.0	0.073	18
16. AM/AM	9600	2300	0.24	9.5

* Here information efficiency is defined as the ratio of the information capacity of the output signal channel to the information capacity of the modulated signal.

This table is based on an individual channel signal-to-noise ratio of 100 to 1 and a total information bandwidth of 1000 cps.

Of the sixteen telemetry systems listed in the table those using PCM stand out above all others except for PPM/AM. PCM is one of the practical developments that has been made in more recent years and is one that coincides closely with theoretical predictions of information theory. While other more recent systems have been used which come closer to realizing theoretical transmission efficiencies[72], PCM is more highly developed and state-of-the-art equipment has met the requirements of the space age.

In addition to the fact that PCM systems exhibit good communication efficiencies, they are digital systems and have all the inherent advantages of digital circuitry. Digital circuits, once properly designed, do not need readjustment.

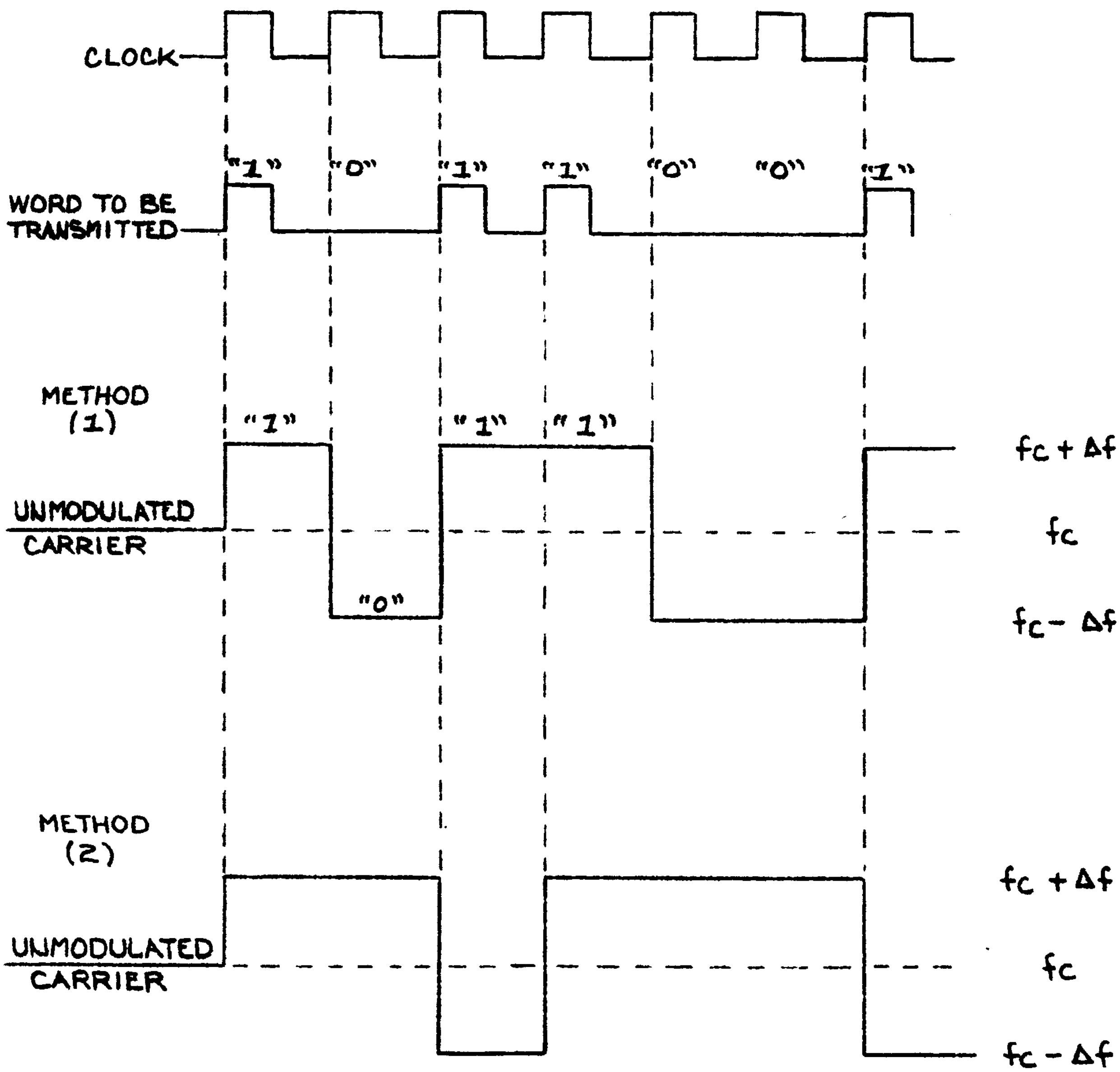
Accuracy and repeatability of data are excellent in a digital system of this type. State-of-the-art PCM equipment for airborne application consistently produces 0.1% accuracy under conditions of vibration, acceleration and temperature environment extremes. This accuracy represents at least a 10-to-1 improvement over the best analog systems.

Time-division multiplexing with PCM systems is extremely flexible, allowing both super and sub-commutation routines to be established at the discretion of the experimenter, in accordance with the data sampling rates required. Systems have been built with approximately 100 channels of data operating with full scale sensitivities as low as 10 millivolts.

Communication using a binary coded digital signal has many advantages over a system in which analog signals are to be transmitted. In the binary coded digital system only two levels are used for encoding all information.

Therefore, all circuits in the communication system are designed to operate at two distinct levels, resulting in considerable simplification of circuit design problems throughout the system. Signal-to-noise ratios in the digital system may be much lower than in an analog system of equal accuracy. The PCM system, in fact, has an overall threshold different from the analog system. In most analog systems once above the detection threshold increasing the signal signal-to-noise ratio at the input to the receiver will increase the signal to noise ratio at the output in a linear manner. With the PCM system a decision circuit follows the detector which has a threshold effect different from the detector. The threshold in the decision circuit is reached when increasing the signal-to-noise ratio at the input has little or no effect on the errors or output signal-to-noise ratio. This threshold is in the vicinity of 15 to 20 DB. The result of these two thresholds in the PCM system is that once above a certain minimum signal-to-noise ratio, essentially error free reception is possible. The only noise that can be considered beyond the decision circuit is the granulation of the signal due to the fact it had to be quantizied in order to code it for transmission. If the receiver signal is to be retained in its digital form, this quantization noise is not a factor for consideration. This fact will become more clear in later analysis.

The basic PCM time division multiplex can be used with several different types of RF modulation including AM, PM, and FM. Of these, either FM or PM is the logical choice. The modulator required for either FM or PM can be smaller, hence lighter weight, and consume less satellite power than the AM type. In addition, the natural suppression of atmospheric noise and adjacent channel interference by the limiting action of the receiver circuits make FM and PM more attractive for reliable communication systems.



Methods of NRZ Modulation
Figure VII.

In any system that has two thresholding effects the most efficient operation is obtained when the system parameters are adjusted so that both thresholds are overcome simultaneously. In the PCM/FM system the parameter can be adjusted to optimize the system is β , the FM deviation ratio defined as $\beta \triangleq \frac{\Delta f}{f_m}$

where Δf = Maximum frequency deviation of the RF carrier

f_m = Maximum modulating frequency.

The method of optimization and the results obtained are illustrated in the analysis that follows.

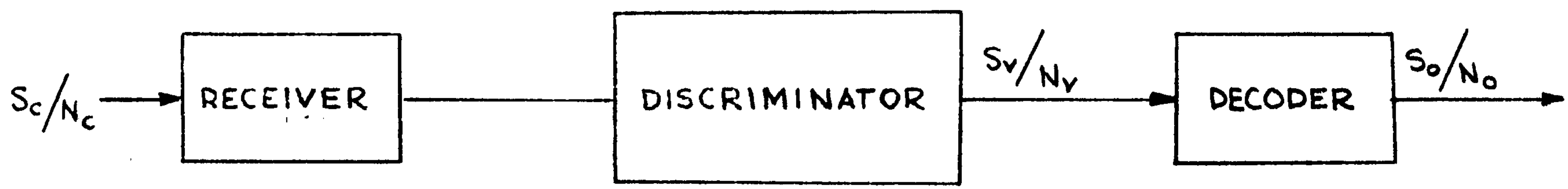
Analysis of Optimum Deviation Ratio

The optimum deviation ratio and hence the receiver prediction bandwidth may be determined by a systematic method of matching signal-to-noise ratios in the receiver system to achieve simultaneous thresholds in the FM discriminator and the PCM decoder. Referring to the block diagram of Figure VIII, the signal-to-noise ratio at the output of the PCM decoder, S_o/N_o, may be interpreted differently depending on how the signal is to be used. If the output is to be an analog signal then S_o/N_o will contain noise due to quantizing the signal and an equivalent noise (error) because of missing (or wrong) bits in the digital signal. When the output is an analog signal that results from decoding a series of digital words the expression for the output signal-to-noise power ratio caused by a finite input signal-to-noise ratio is:[13]

$$\frac{P_o}{N_o} = \left(\frac{S_o}{N_o}\right)^2 = \frac{3\sqrt{3}}{8} \cdot m^2 \cdot \frac{2^{\mu} - 1}{2^{\mu} + 1} \cdot \left(\frac{\mu}{f_o \mathcal{T}}\right) e^{\frac{r^2}{2}}$$

where:

$$r^2 = \left(\frac{S_v}{N_v}\right)^2 = \text{Power ratio at the input}$$



Block Diagram - Receiver System
Figure VIII.

T = Duration of a coded word

μ = Number of code pulses in T

f_o = Cutoff frequency of the output filter = $\frac{0.5}{T}$ for a typical matched filter case.

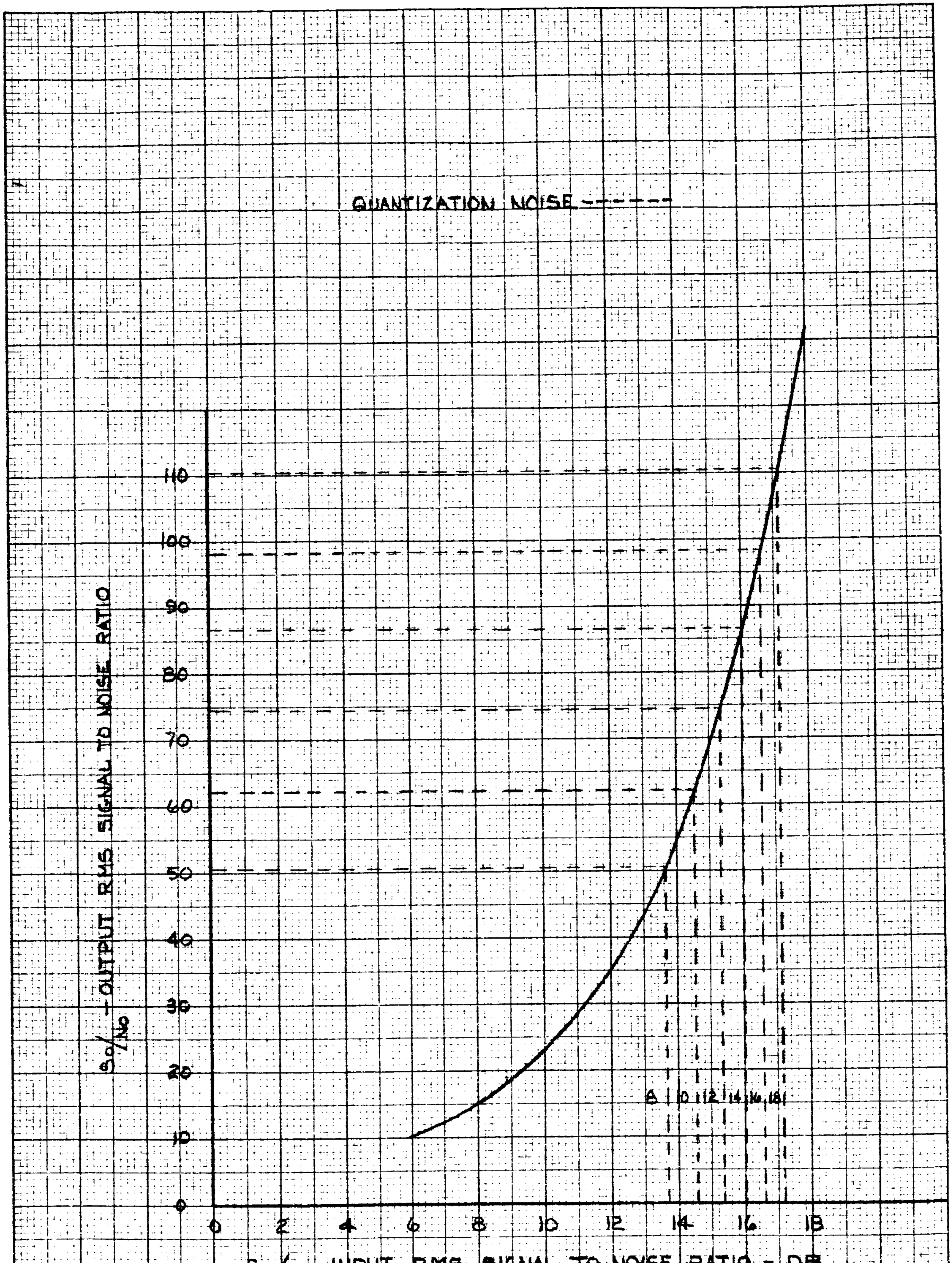
m = Degree of modulation, assumed = 1.0

The equation is plotted in Figure IX. Also noted on the graph is the noise at the output that results from the granularity of sampling and quantizing a signal. This noise level is given by:

$$S_o/N_o = \sqrt{3/2} \; : 2^{\mu}$$

This latter expression for noise is based on the assumption that the input signal-to-noise ratio to the decoder is well above the PCM threshold and the only contribution to the output noise is just that caused by the minimum quantizing increment. If an analog output signal is desired from the PCM system then the input signal-to-noise ratio for most efficient operation is that signal-to-noise input ratio that makes the output signal-to-noise ratio just equal to that which results from quantization noise alone. The total output noise will then be the sum of that caused by the input noise and the quantization noise or twice the noise power that would result from either effect alone. The resultant output signal-to-noise ratio should therefore be reduced 3 DB.

In the PCM system being considered here the decoder serves only to make the decision as to whether a "one" or a "zero" has been received. The output is thus the binary coded digital word that was transmitted, if no decision errors are made. In this case noise as such does not have meaning. Instead it is customary to measure the quality of the output signal in terms of the



Signal-to-Noise Improvement in a PCM Decoder
Figure IX.

number of right and wrong code pulses that have been received, hence error rate or bits-without-error is a measure of the quality of digital transmission when the information remains in digital form.

To provide quanitative answers to the question of what noise level is tolerable in a PCM decoder circuit it is necessary to analyze in greater detail the mechanism for deciding whether a "one" or a "zero" is present in the incoming digital code.

Assume first, a binary type on-off PCM system in which pulses represent "1's" and no pulses represent "0's". In the selection bandpass system of the receiver fluctuation noise is added to the incoming group of RF pulses. The RF signal is then demodulated and presented to the PCM decoder. This fluctuation causes error in the decision ability of the decoder. An error will occur in the absence of a pulse if the amplitude of the noise, at the instant that the decision is made, is greater than the decision level. An error will be made in the presence of a pulse signal if the noise is of opposite polarity and of sufficient amplitude to drive the signal below the decision level.

In the first case the decoder would have decided a "1" was present when actually a "0" was transmitted and in the second case the decoder would have sensed a "0" when actually a "1" was present.

Assume that the fluctuation noise at the input to the decoder is Gaussian with mean value equal to zero and that the amplitude of the pulse signal is A_c volts. The decision level in the PCM decoder is set at $A_c/2$ volts. Assume first a "0" is sent so that no pulse exists at the input to the decoder. The probability that an error in decision will be made is just the probability that the amplitude of the noise will exceed $+ A_c/2$ volts and be mistaken for a "1". If the voltage at the input to the decoder is designated as v (t) then the probability of an error is just the probability that v (t) > $A_c/2$ which is given by the area under the probability density curve from $A_c/2$ to ∞ ,

$$P(\text{"0" error}) = P(v(t) > A_c/2) = \int_{A_c/2}^{\infty} \frac{e^{\frac{-v^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, dv$$

where, σ^2 = (rms noise voltage)^2

Pictorially this is represented by the shaded area in Figure Xa.

Assume now that a "1" is transmitted. It will appear at the input to the decoder as A_c volts plus superimposed fluctuation noise. The signal plus noise will have a mean value of A_c volts and the probability density function is given by

$$p(v) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\frac{-(v - A_c)^2}{2\sigma^2}}$$

The signal plus noise will be higher than the clipping level, $A_c/2$ unless the noise is negative and of an amplitude greater than the $A_c/2$ such as to reduce the total signal to less than the decision level.

The probability of error is thus just the area of the probability density function that lies below $A_c/2$ volts is given by,

$$P(v < A_c/2) = \int_{-\infty}^{A_c/2} \frac{e^{-\frac{(v - A_c)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, dv$$

This probability is illustrated by the shaded area in Figure Xb.

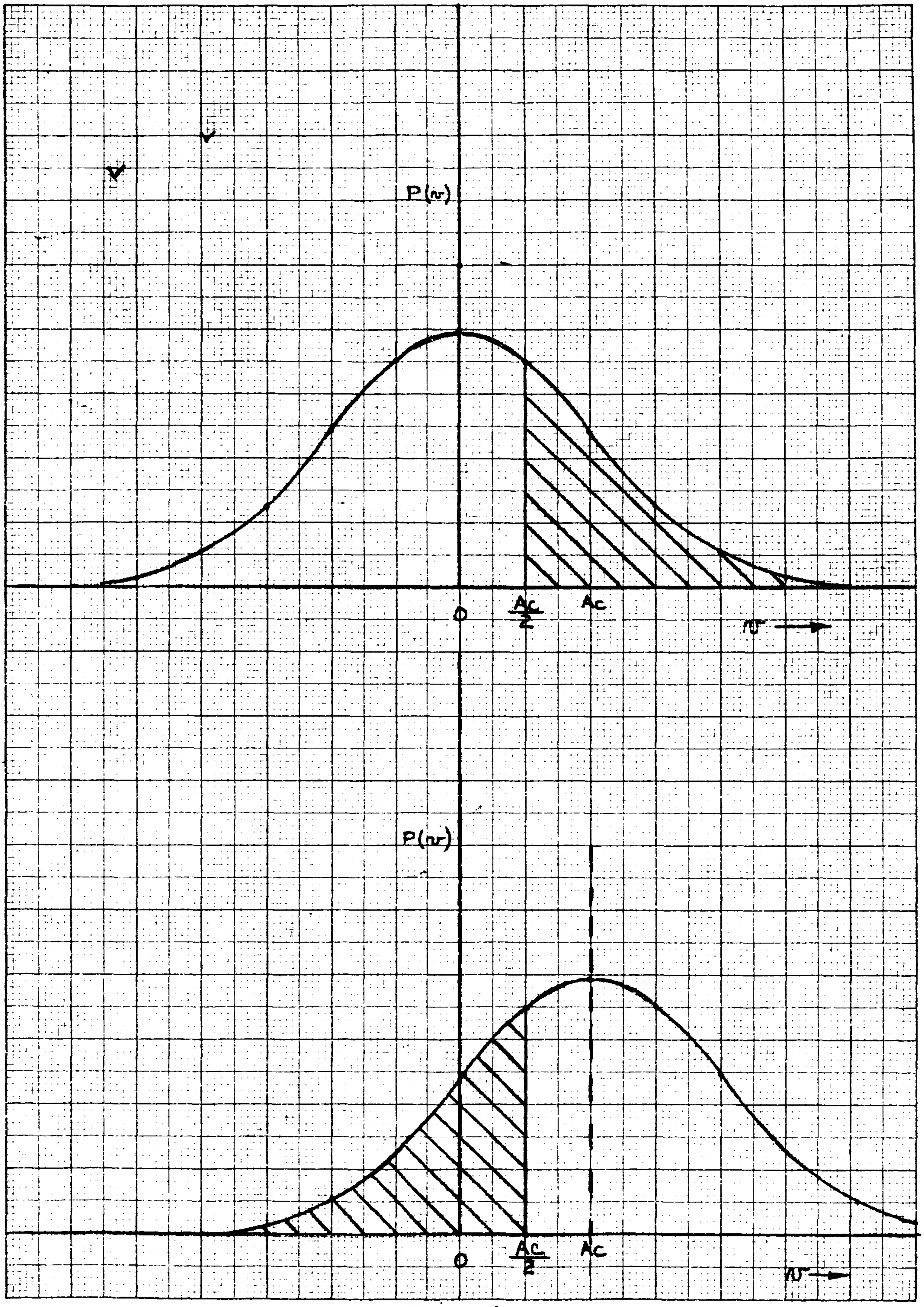


Figure X.
Normal Distribution Curve

This probability is the same as the probability that noise alone will be more negative than - $A_c/2$ volts, as is illustrated by letting $X = v - A_c$

$$P(\text{"1" error}) = P(v < A_c/2) = \int_{-\infty}^{-A_c/2} \frac{e^{-\frac{X^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, dx$$

and it can be seen that the two areas are equivalent, and that the probability of mistaking a "0" for a "1" is equal to the probability of mistaking a "1" for a "0" if the noise is random with Gaussian Distribution and mean zero.

If the probability of a "1" or a "0" appearing in the message is equally likely, P (1) = P (0) = .5, as is the general case for data transmission, then setting P ("1" error) = P ("0" error), the total probability of error becomes

$$P(\text{error}) = P(\text{"1" error}) = P(\text{"0" error})$$

If the bipolarity pulses are sent as recommended by IRIG standards, their amplitude can be - $A_c/2$ and + $A_c/2$. The decoder then must decide whether the instantaneous sum of the signal plus noise is positive or negative. The probability of a decision error for a "1" transmitted (assuming a "1" is sent as + $A_c/2$

$$P(\text{"1" error}) = \int_{-\infty}^{0} \frac{e^{-\frac{(v - A_c/2)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, dv = \int_{-\infty}^{A_c/2} \frac{e^{-\frac{v^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

which is the same as the case for the undirectional pulse of amplitude A_c. Similarly the probability of error for a bipolar "0" of amplitude - $A_c/2$ is the same as the probability previously derived. Curves drawn for the probability of error versus signal to noise ratio will apply for either case if the proper abscissa is chosen.

For a total voltage swing of A_c volts into the PCM decoder it is necessary to know the average power and rms voltage in the signal waveform. With this information the probability curves can be plotted in the most useful form.

In the binary coded pulses two possibilities are evident:

(1) With on-off signals, the on signal is a pulse of amplitude A_c volts. The decoder decision level is set to make the decision at $A_c/2$ volts. The peak power of the signal is A_c^2 if it is assumed that this power is dissipated in a one ohm load. If the signals "1" and "0" are equally likely then the average power is just one-half the peak power or

$$P_s \text{ (avg)} = \frac{P_s(\text{Peak})}{2} = \frac{A_c^2}{2}$$

The rms voltage is,

$$S \text{ (rms)} = \sqrt{P_s \text{ (avg)}} = \frac{A_c}{\sqrt{2}}$$

(2) If plus and minus pulses of amplitude $+ A_c/2$ and $- A_c/2$ are sent instead, the decoder makes the decision on the polarity of the instantaneous voltage.

Assume again that "1's" and "0's" are equally likely, then the peak power in a one ohm resistor is,

$$P_s \text{ (peak)} = \left(\pm A_c/2\right)^2 = \frac{A_c^2}{4} = P_s \text{ (avg)}$$

Comparing the results, less power is required for the plus-minus transmission with the same probability of error out of the decoder. Hence, this is the preferred method of transmission.

If it is equally likely that "1's" and "0's" will be transmitted then,

$$\frac{\sum \text{error}}{\text{total bits}} = P\ (\text{"0" error}) = \int_{A_c/2}^{\infty} \frac{e^{-\frac{v^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, dv$$

which may be written as,

$$\frac{\text{errors}}{\text{total bits}} = 1 - \int_{-\infty}^{A_c/2} \frac{e^{-\frac{1}{2}\left(\frac{v}{\sigma}\right)^2}}{\sqrt{2\pi\sigma^2}}\, dv = Q\ (Z) = 1 - F\ (Z)$$

$Z = \frac{v}{\sigma}$

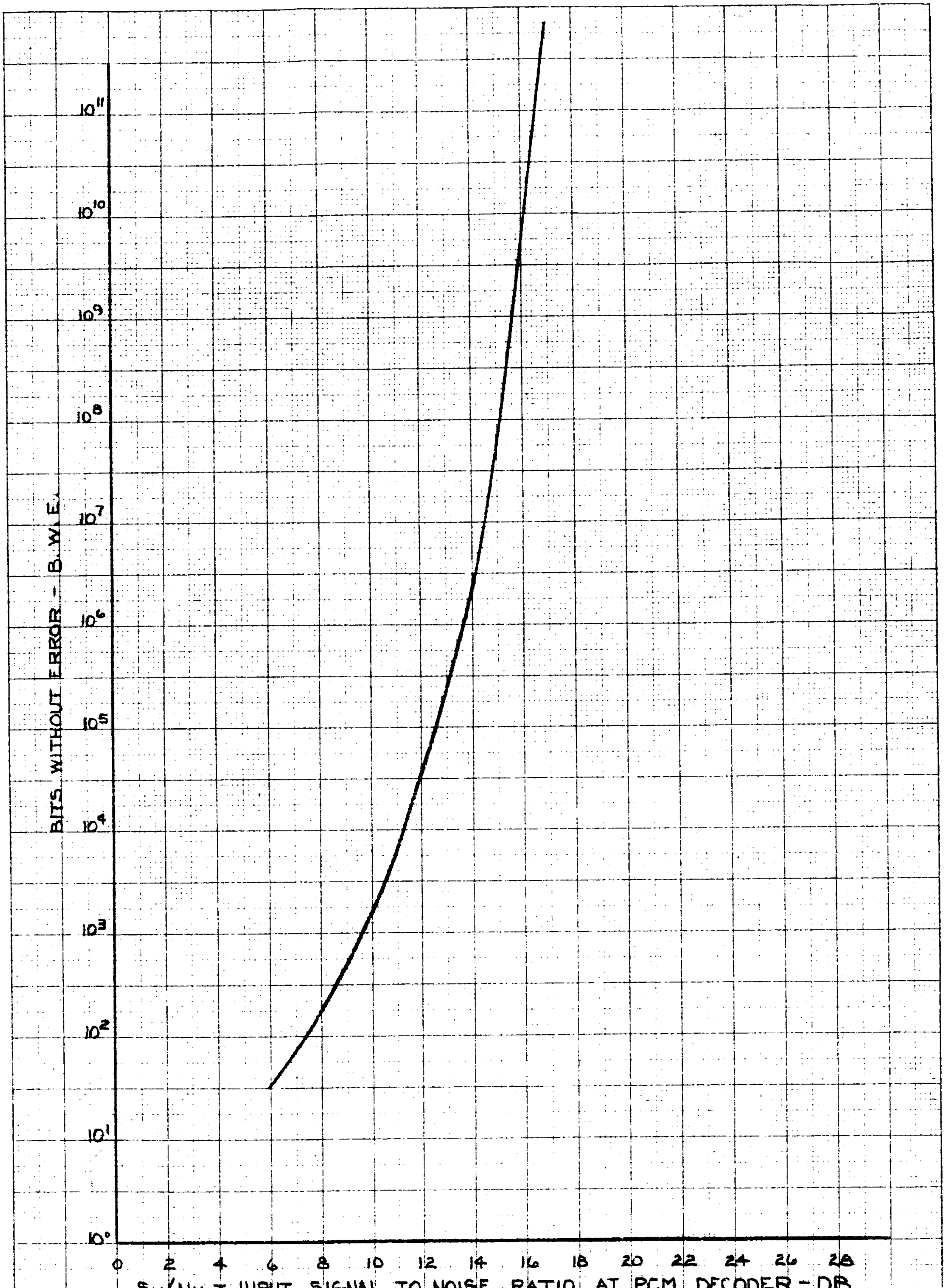
$F\ (Z)$ = percentage of time v lies above $A_c/2$

$Q\ (Z)$ = percentage of time v lies below $A_c/2$

$Z_1 = A_c/2$ = rms signal-to-noise ratio.

Values of Q (Z) are tabulated in tables of the Normal Distribution Function. The reciprocal of Q (Z) is the desired quantity, bits-without-error (BWE) and, is shown plotted in Figure XI where the notation S_{V/N_V} has replaced $A_c/2\sigma$ for the input rms signal-to-noise ratio.

Of interest here is the so-called "PCM threshold." For the case shown the threshold might be defined as being in the vicinity of 12 DB input signal-to-noise ratio. If the signal-to-noise ratio falls by 3 DB the bits-without-error decreases by 100 to 1. If the input signal level is raised by 3 DB the number of bits without error increases by nearly 1000 to 1. A signficant decrease in error rate for a small signal level change. However, once above approximately 15 DB signal-to-noise ratio the error rate is so small as to be nearly insignificant; thus the so-called PCM threshold.



Bits Without Error vs. Input Signal-to-Noise Ratio
Figure XI.

As a design parameter a bits-without-error (BWE) of 10^9 will be chosen requiring an input S_{v/N_v} to the decoder of 15.6 DB.

Having chosen the minimum signal-to-noise ratio at the input to the PCM decoder (output of the FM discriminator) it is now possible to examine the affect that the parameters of the FM system have on this discriminator output signal-to-noise ratio.

In order to write an expression for a frequency modulated signal, it is necessary to change the concept of frequency slightly. Normally we write for a cosine wave of frequency ω_c,

$$f_c(t) = \cos\Theta(t) = \cos(\omega_c t + \Theta_o)$$

wherein $\Theta(t)$ is a linear time variant with constant derivative ω_c. However, for FM, $\Theta(t)$ is not a linear function of time and hence its derivative is not a constant but a function of the modulating signal. To eliminate this difficulty and still maintain the concept of frequency, define an instantaneous angular frequency ω_i such that

$$\omega_i \triangleq \frac{d\left[\Theta(t)\right]}{dt}$$

Let

$$\Theta(t) = \omega_c t + \Theta_o + k_1 f(t)$$

where f(t) is the modulating signal.

If K_1 is constant, this is phase modulation since the phase of $f_c(t)$ varies linearly with the modulating signal. Now let the frequency vary as a function of time and such that

$$\frac{d\,\Theta(t)}{dt} = \omega_i = \omega_c + k_2 f(t)$$

then

$$\Theta(t) = \int \omega_i\, dt = \omega_c t + \Theta_o + k_2 \int f(t)\, dt$$

and the phase of $f_c(t)$ varies as the integral of the modulating signal. The FM signal of interest is $f_c(t)$ when (t)

$$(t) = {}_c t + {}_o + k_2 \quad f(t)\, dt$$

For a simple analysis let the modulating signal be a sinusoid with frequency

$$\omega_m = 2\pi f_m,$$

$$f(t) = a \cos \omega_m t$$

then

$$\omega_i = \omega_c + \Delta\omega \cos \omega_m t \quad \text{where} \quad \Delta\omega = 2\pi\Delta f,$$

Δf being the maximum frequency excursion set by the circuitry of the system. Then

$$\Theta(t) = \int \omega_i \, dt = \omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t + \Theta_o$$

If the appropriate phase reference is assumed then the expression for the modulated carrier can be written as

$$f_c(t) = \cos \Theta(t) = \cos\left[\omega_c t + \beta \sin \omega_m t\right]$$

where

$$\beta \triangleq \frac{\Delta\omega}{\omega_m} = \frac{\Delta f}{f_m}$$

β is called the modulation index, and is the ratio of the maximum frequency deviation to the modulating frequency. β also is the maximum phase shift of the carrier, as may be seen by reference to the expression for $\Theta(t)$.

Expanding the above expression for $f_c(t)$ by trigonometric identities,

$$f_c(t) = \cos \omega_c t \cos(\beta \sin \omega_m t) - \sin \omega_c t \sin(\beta \sin \omega_m t) .$$

Expanding two of the factors,

$$\cos(\beta \sin \omega_m t) \text{ and } \sin(\beta \sin \omega_m t)$$

into power series,

$$\sin(\beta \sin \omega_m t) = \beta \sin \omega_m t - \frac{\beta^3 \sin^3 \omega_m t}{3} + \frac{\beta^5 \sin^5 \omega_m t}{5} + \cdots$$

$$\cos(\beta \sin \omega_m t) = 1 - \frac{\beta^2 \sin^2 \omega_m t}{} + \frac{\beta^4 \sin^4 \omega_m t}{} + \cdots$$

It is seen from the expansion that each of the terms are periodic function of ω_m. Each contains terms of ω_m and its harmonics. Each harmonic, when multiplied by cos $\omega_c t$ or sin$\omega_c t$, will give rise to two symmetrical side bands about ω_c. The side bands associated with the sine term will be in phase quadrature with the cosine term, the latter being in phase with the carrier.

Of greatest interest is the term cos $(\beta \sin \omega_m t)$. When this term is plotted as a function of ω_m for various values of β some interesting conclusions may be drawn.

Figure XII plots this expression for three values of   and the following discussion covers the pertinent points concerning these plots.

$\beta = 0.5$ and $\beta = 1.0$

For these small values of β the curve varies between 1.0 and .54 at twice the frequency of the modulating waveform. When the function is multiplied by the carrier term cos $\omega_c t$ it will give rise to second order sideband terms $\pm 2\,\omega_m$, $\pm 4\,\omega_m$, etc. about the carrier. The result, then, is that the carrier varies in amplitude at twice the modulating frequency. However, with β less than $\pi/2$ it never goes completely to zero or negative. The variations in the carrier

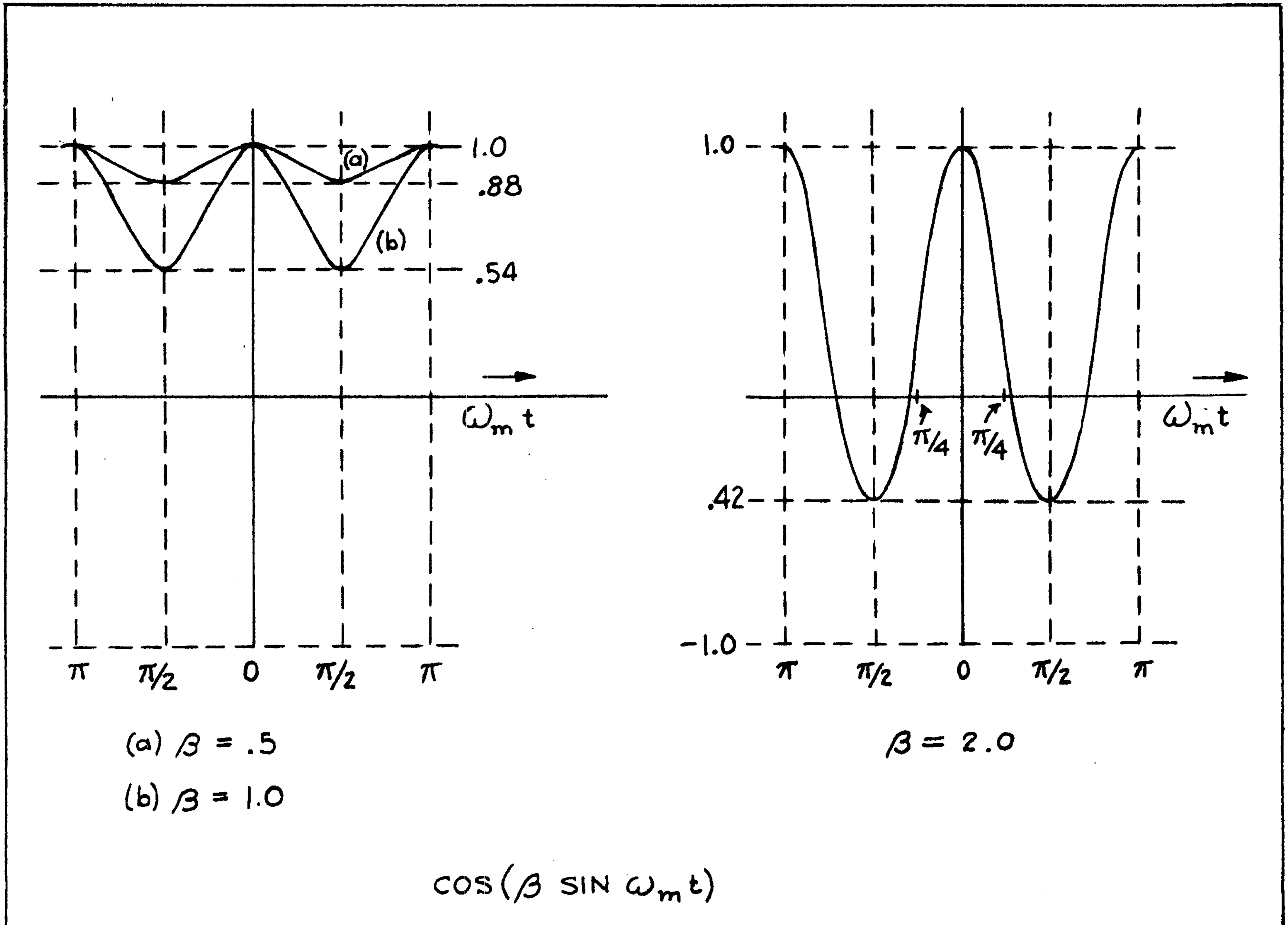

1.0
.88
.54
(a)
(b)
$\omega_m t$
π
π/2
0
π/2
π
(a) β = .5
(b) β = 1.0
1.0
.42
−1.0
π/4
π/4
$\omega_m t$
π
π/2
0
π/2
π
β = 2.0
cos(β sin $\omega_m t$)

Figure XII.

amplitude are larger with larger β. The sidebands created are second order and will increase with larger β.

$\underline{\beta = 2.0}$

For β larger than $\pi/2$, the carrier amplitude varies such as to go to zero and even take on negative values. Since the power in the carrier is less, and the power in a FM signal is constant then more and more of the power must be converted into the sidebands with increasing β, i.e., the spectral energy is spread over a larger region in the frequency domain.

The results from plotting the sine term are similar except that the amplitude varies symmetrically about zero, and its frequency components are all odd integral multiples of the modulating frequency ω_m. When multiplied by $\sin \omega_c t$ they give rise to odd-order sidebands about ω_c.

The amplitude of each of the frequency components of the complex FM wave can be quantitatively determined by expanding the function into a Fourier series. Consider the complex periodic exponential,

$$v(t) = e^{j\beta \sin \omega_m t} = \cos(\beta \sin \omega_m t) + j \sin(\beta \sin \omega_m t)$$

over a period,

$$-T/2 < t < +T/2.$$

The Fourier coefficients for this function will have a real part consisting of even harmonics and an imaginary part consisting of odd harmonics. By equating the reals and imaginaries, the expansion of

$$\cos(\beta \sin \omega_m t)$$

and

$$\sin(\beta \sin \omega_m t)$$

can be obtained simultaneously.

Applying this sort of analysis to the expression determined earlier for the modulated wave,

$$f_c(t) = \cos \omega_c t \cos (\beta \sin \omega_m t) - \sin \omega_c t \sin (\beta \sin \omega_m t)$$

after a great deal of rather complex analysis[9], the following expression results:

$$f_c(t) = J_o (\beta) \cos \omega_c t + \sum_{n=1}^{\infty} J_n (\beta) \quad \cos (\omega_c + n\omega_m) t + (-1)^n \cos (\omega_c - n\omega_m) t$$

where $J_n (\beta)$ are bessel functions of the first kind.

The complete FM signal then consists of a carrier term plus upper and lower sideband terms displaced from the carrier by integral multiples of the frequency of the modulating signal. The amplitude of these sideband terms, as controlled by multiplier $J_n (\beta)$, can be examined to determine how they vary with values of β .

Table II tabulates the values of $J_n (\beta)$ for various values of β . Also noted in the table is the RF bandwidth required to pass the FM signal when all sidebands over 1% are necessary and when only the sidebands greater than 10% are necessary. For the telemetry of analog signals, where non-linearities in the data link can affect the resultant data, the 1% sidebands are necessary. However, in a PCM digital data link, the 10% sidebands are considered adequate for "1" or "0" recognition.

Figure XIII plots the required RF bandwidth to pass all sidebands greater than 10% versus deviation ratio. The RF bandwidth, (BW = 2nfm) is normalized by the factor $2f_m$ such that the plot is actually just the number of significant sidebands. The plot has been made continuous, rather than discrete as tabulated, for later usefulness. It can be noted that above $\beta = 1$ the

TABLE II

β	$J_0(\beta)$	$J_1(\beta)$	$J_2(\beta)$	$J_3(\beta)$	$J_4(\beta)$	No. of Significant Sidebands $>1.0\%$	$>10\%$	Bandwidth Required 1% Side-bands	10% Side-bands
0.1	.9975	.0499	.0012	.0000	.0000	1	0	2fm	2fm
0.2	.9900	.0995	.005	.0002	.0000	1	0	2fm	2fm
0.3	.9776	.1483	.0112	.0006	.0000	2	1	4fm	2fm
0.4	.9604	.1960	.0197	.0013	.0001	2	1	4fm	2fm
0.5	.9385	.2423	.0306	.0026	.0002	2	1	4fm	2fm
0.6	.9120	.2867	.0437	.0044	.0003	2	1	4fm	2fm
0.7	.8812	.3290	.0588	.0069	.0006	2	1	4fm	2fm
0.8	.8463	.3688	.0758	.0102	.0010	3	1	6fm	2fm
0.9	.8075	.4059	.0946	.0144	.0016	3	1	6fm	2fm
1.0	.7652	.4401	.1149	.0196	.0025	3	2	6fm	4fm
1.2	.6711	.4983	.1593	.0329	.0050	3	2	6fm	4fm
1.4	.5669	.5419	.2074	.0505	.0091	3	2	6fm	4fm
1.6	.4554	.5699	.2570	.0725	.0150	4	2	8fm	4fm
1.8	.3400	.5815	.3061	.0988	.0232	4	2	8fm	4fm
2.0	.2239	.5767	.3528	0.1289	.0340	4	3	8fm	6fm
2.5	(-).0484	.4971	.4461	.2166	.0738	4	3	8fm	6fm
3.0	(-).2601	.3391	.4861	.3091	.1320	>4	4	>8fm	8fm
4.0	(-).3971	(-).0660	.3641	.4302	.2811	>4	>4	>8fm	>8fm

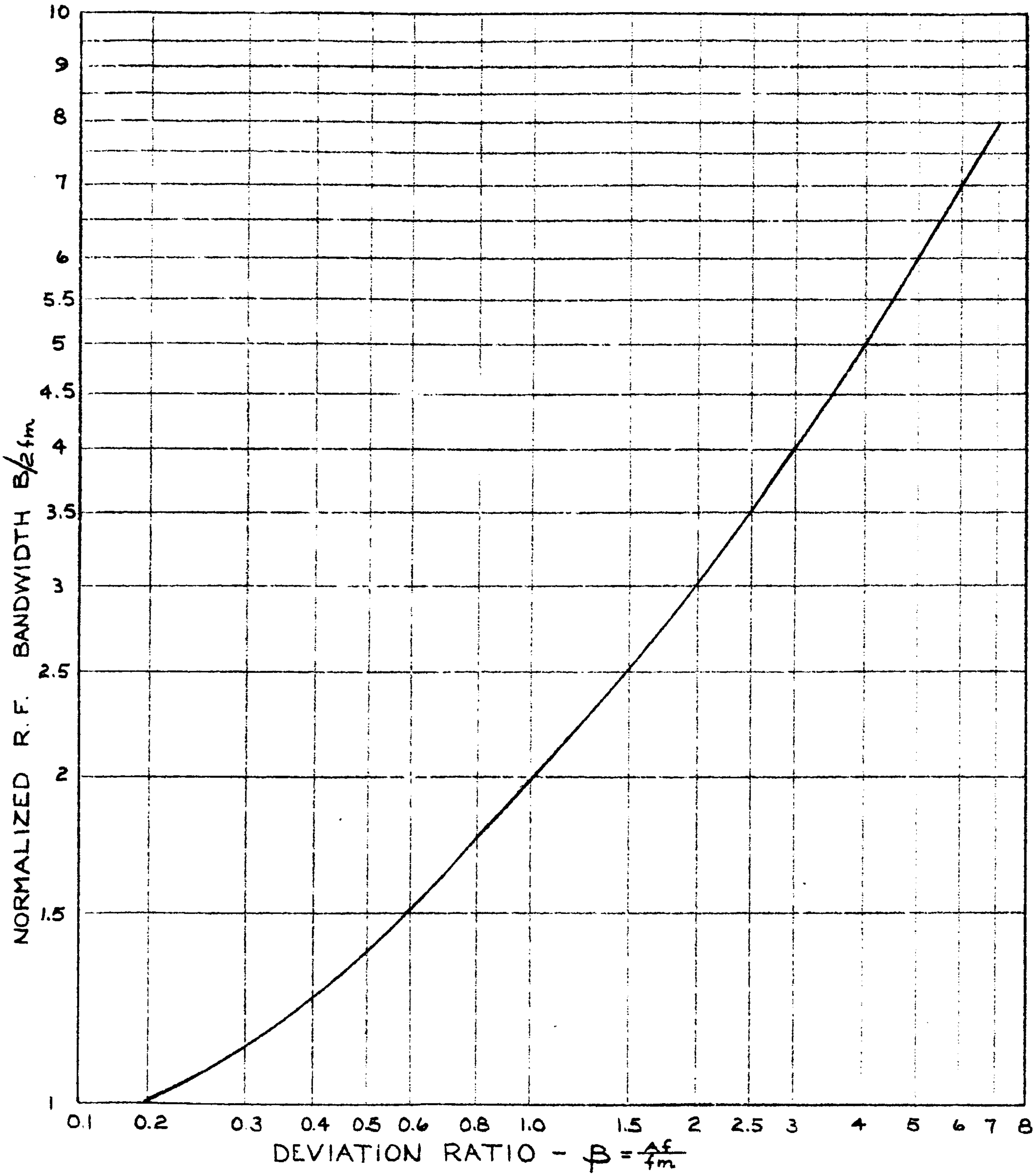


Figure XIII.
RF Bandwidth vs. Deviation Ratio

number of sidebands required, n, is $\beta + 1$. Where β is not an integer the curve is continuous and approximately maintains the same relationship.

The pre-detection bandwidth of the receiver is therefore given approximately by

$$B \approx 2\,(\beta + 1)\,f_m$$

Crosby's[19] original work on the characteristics of FM in the presence of noise yield several relationships that may be used here.

(1) The signal-to-noise power ratio at the output of an FM discriminator and low pass filter with cutoff frequency f_m is given by:

$$\frac{P_{sv}}{P_{nv}} = 3\beta^2 \, \frac{P_{sc}}{P_{nc}} \text{ (AM)} \quad \text{where}$$

$\dfrac{P_{sv}}{P_{nv}} = \left(\dfrac{S_v}{N_v}\right)^2$ = signal-to-noise power ratio at the output of the FM discriminator

$\dfrac{P_{sc}}{P_{nc}}$ (AM) = carrier-to-noise power ratio in the pre-detection bandwidth $2\,f_m$ of an AM system = $\left(\dfrac{S_c}{N_c}\right)^2$ (AM)

β = deviation ratio

(2) The threshold in an FM system is defined as the signal-to-noise ratio above which the full FM improvement given in (1) above is realized and below which the output signal-to-noise ratio falls rapidly below that of an AM system. This threshold may be approximated by a point where the input peak signal to peak noise voltage ratio is 1.414.

(3) The crest factor of noise can be taken as being approximately 12 DB = 4.0

With these relationships and the previous discussions of the bandwidth requirements of an FM signal it is now possible to establish the deviation ratio that will set the FM threshold and the PCM threshold so that both occur at the same carrier-to-noise ratio.

At the FM threshold, the peak signal to peak noise voltage ratio is given by:

$$\frac{S_c\ (\text{peak})}{N_c\ (\text{peak})}\ (\text{FM}) = 1.414 = \frac{S_c\ (\text{rms})}{N_c\ (\text{rms})} \cdot \frac{1.414}{4.0}$$

solving for the rms signal-to-noise ratio in the FM pre-detection bandwidth

$\frac{S_c}{N_c} = 4.0$, and the power ratio is,

$$\frac{P_{sc}}{P_{nc}}\ (\text{FM}) = \left(\frac{S_c}{N_c}\right)^2 = 16.0$$

The above carrier-to-noise ratio for an FM system may be readjusted for the AM case by multiplying by the ratio of the predetection bandwidth.

$$\frac{P_{sc}}{P_{nc}}\ (\text{AM}) = \frac{P_{sc}}{P_{nc}}\ (\text{FM}) \times \frac{B(\text{FM})}{B(\text{AM})} = \frac{P_{sc}}{P_{nc}}\ (\text{FM}) \cdot \frac{2f_m\ (\beta + 1)}{2f_m}$$

$$= \frac{P_{sc}}{P_{nc}}\ (\text{FM})\ (\beta + 1)$$

Substituting this expression for the AM carrier to noise power ratio in the equation for the output signal-to-noise ratio, we get

$$\frac{P_{sv}}{P_{nv}} = 3\beta^2 \frac{P_{sc}}{P_{nc}}\ (\text{AM}) = 3\beta^2 \frac{P_{sc}}{P_{nc}}\ (\text{FM})\ (\beta + 1)$$

This equation may now be solved for β if the power ratio at the input to the decoder, $\frac{P_{sv}}{P_{nv}}$ is set equal to the PCM threshold value of 15.6DB = 36.2 and $\frac{P_{sc}}{P_{nc}}$ (FM) is set to its threshold value of 16.

Then

$$\beta^2 (\beta + 1) = \frac{36.2}{(3)\ (16)} = .755$$

The positive solution for β is the desired one and solving the equation gives β = .672.

This value of deviation ratio yields an output signal-to-noise ratio from the discriminator of 15.6DB at the FM threshold. This is equal to the minimum signal-to-noise ratio desired at the input to the PCM decoder. These values are achieved with an FM carrier-to-noise power ratio of 16 = 12DB.

By combining the results of the previous sections, it is now possible to determine the required transmitter power as a function of the bit rate that is to be transmitted.

The block diagram of the general communication system link being considered is shown in Figure XIV, with the gains of each of the blocks labeled.

From this figure the following relationships may be derived,

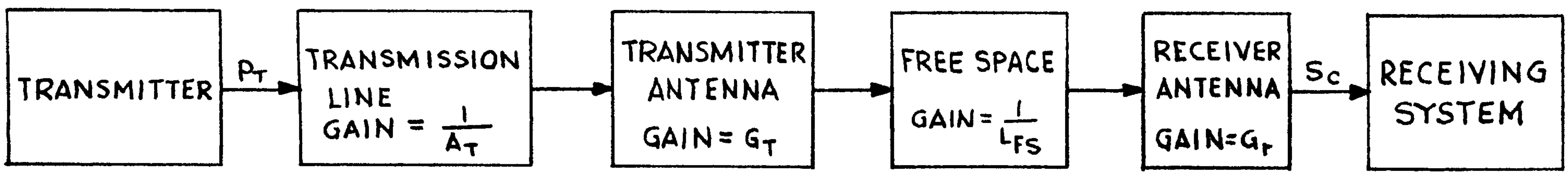
$P_{sc} = P_t\, G_{total}$, **where:** G_{total} **= overall link gain =** $G_1 \cdot G_2 \cdot G_3$ **etc.**

$$\frac{P_{sc}}{P_{nc}} = \frac{P_t\, G_{total}}{P_{nc}}$$ **= signal-to-noise ratio at the input to the receiver**

The required transmitter power for a particular signal to noise ratio $\frac{P_{sc}}{P_{nc}}$ at the input to the receiver is,

$$P_t = \frac{P_{sc}}{P_{nc}} \ \frac{P_{nc}}{G_{total}}$$

where $$G_{total} = \frac{G_t\, G_r}{A_r L_{fs}}$$



General Communication System
Figure XIV

and

A_T = Transmission line loss at the transmitter = 1.26 = 1.0 DB

L_{fs} = Free space loss $= \left(\frac{4\pi d}{\lambda}\right)^2$

d = Path distance

G_T = Transmitter Antenna Gain

G_r = Receiving Antenna Gain = $.54 \left(\frac{\pi D}{\lambda}\right)^2$ for parabolic antenna of diameter D.

λ = Wavelength

P_{nc} = Effective noise referred to the input terminals of the receiving system.

$= K B (T_{eff} + T_o)$

K = Boltzmann's Constant = 1.38×10^{-23} joules per ^oK

B = Pre-detection Bandwidth = $2f (\beta + 1)$

T_{eff} = Effective temperature of the receiving system

f_m = Maximum modulating frequency = $\frac{\text{max bit rate}}{2} = \frac{B_r}{2}$ using NRZ coding

In terms of these values P_t becomes

$$P_t = \frac{P_{sc}}{P_{nc}} KB_r (\beta + 1) \frac{1.26\ L_{fs}}{G_r} (T_{eff} + T_o)$$

From previous sections of this report the following values for the parameters in this equation have been either plotted graphically or derived,

β = .672

$\frac{P_{sc}}{P_{nc}}$ = 36.2

G_r = 30 DB = 10^3 (assumed value)

L_{fs} = 151 DB = 1.26×10^{15} = free space loss at maximum transmitting distance of 2060 miles (See Figure III).

Substituting these values,

$$P_t = B_r \ (T_{eff} + T_o) \times 13.27 \times 10^{-10}$$

The effective receiver temperature, T_{eff} may be taken from the graph plotted in Figure II for the particular operating frequency. Figure XV plots the required transmitter power in DBM vs bit rate for three carrier frequencies of interest.

<u>Sample Calculation:</u>

Assume: (a) Carrier Frequency = 275 MC

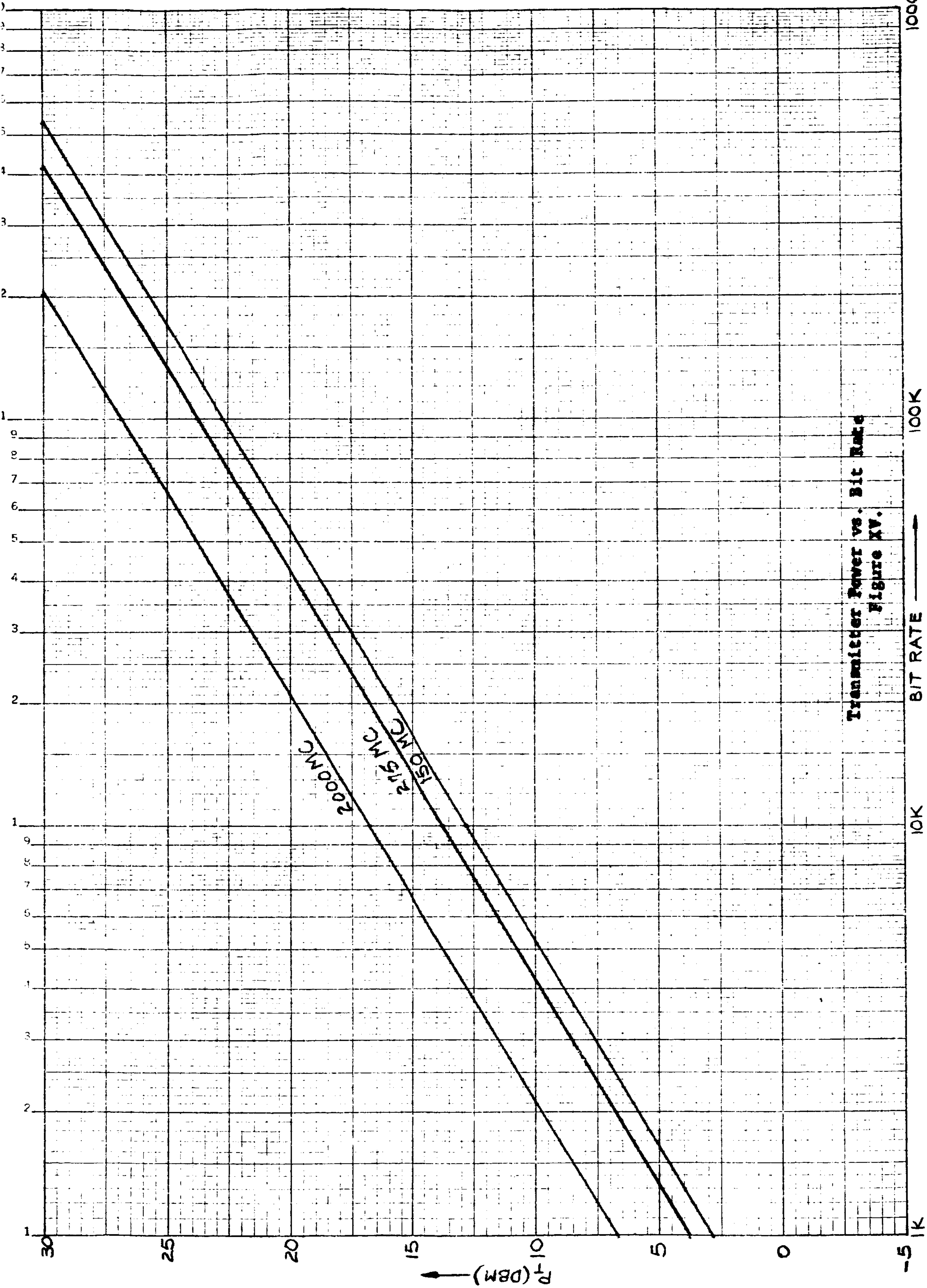
(b) Bit rate = 10 KC = 10^4

$$T_{eff} = 1550^\circ \text{ K}$$

$$P_t = 1840 \times 10^4 \times 13.27 \times 10^{-10} = 24.4 \times 10^{-3} \text{ watts}$$

$$P_t \text{ (DBM)} = 13.86 \text{ DBM}$$

The results of the graph plotted indicate the minimum RF power required at the output of the transmitter to maintain good error free reception of PCM data with the assumed system parameters. It should be noted that no allowance has been made for the fading typical of long range transmission paths. It would seem reasonable to allow as a minimum 3DB (50%) tolerance for deterioration of transmitter power output with age. The fading allowance of 18 DB as shown in the section of the report on RF characteristics would then raise the required transmitter output power by 21 DB over the levels plotted.



Transmitter Power vs. Bit Rate
Figure XV.

V. CODING METHODS AND AIMS

In the previous section on modulation and multiplexing, it was concluded that there are distinct advantages in using a pulse coded trans mission system for achieving a desired receiver system signal-to-noise ratio with a given transmitter power. The problem of the actual coding methods to be used must also be examined. As a starting point, the following general statements may be made.

1. All other parameters being constant, the transmitter power required to achieve a given system output signal-to-noise ratio is proportional to the bandwidth of the information channel. In other words, it must be assumed in the case of PCM systems that the required transmitter power is directly proportional to the maximum bit rate of the system, and vice versa. Based on a consideration of transmitter power alone, one may conclude the goal of coding should be to reduce to a minimum the total average number of bits required to transmit the various possible messages. Note that statistical considerations are implied in the word "average". So-called minimum redundancy coding methods have been developed to achieve this end, where advantage is taken of the statistical properties of the messages to be transmitted. Implied in this technique is the requirement for temporary storage of sequences of messages, so that short time-variations in message source rate may be averaged to achieve a lower over-all transmission rate.*

2. For a given input signal-to-noise ratio, there is always a non-zero error probability. Since, however, the probability decreases exponentially with the transmitter power, almost any desired error rate can be achieved by increasing transmitter power. In a practical case, however, the transmitter

*The term "message" as used here implies one of a set of symbols or groups of symbols designating some quantity or state to be transmitted. In general, messages are not statistically independent. However, only coding of independent messages will be treated here.

power is established by signal-to-noise and power supply considerations of the overall system. Thus the average system error rate may be satisfactory for most of the messages transmitted, but may not be low enough for some particular class of messages which, for various reasons, must be transmitted more reliably.

In general, the error probability can be significantly reduced by adding carefully selected redundancies to the code, so as to permit detection and correction of errors. Note that this technique tends to oppose the goal of decreasing transmission bandwidth.

Simple Redundancy Reducing Codes for Indpendent Messages

Examples of redundancy reducing codes date back at least to the Morse telegraph code. The general approach is to assign shorter code words to messages or symbols occurring most frequently. Intuitively, this should reduce the average number of code pulses transmitted over a long period of time. In constructing a useful and efficient code, careful consideration of the statistical properties of the messages being coded must be made. It is not the purpose of this discussion to explore the underlying theory of optimum coding, but rather to list some of the more useful relationships and methods resulting from it, as they apply to some simple coding problems.* The basic premise is that of Shannon's information theory, i.e., the information is related to the probability of certain messages being communicated. In this theory, information is only measured in terms of the statistics of messages, and nothing is said about meaning.

Entropy as a Measure of the Information Content of a Set of Discrete Messages

Consider a set of N statistically independent messages (m_i) with

*For a more detailed discussion, the reader is referred to texts on the Statistical Information Theory such as: Shannon and Weaver, The Mathematical Theory of Communication; Feinstein, Foundation of Information Theory; Black, Modulation Theory.

probabilities p_i. We can argue intuitively that $I_i = \log (1/_{pi})$ is a logical measure of the information contained in the selection of messages (m_i). It follows the notion that if $p_i \approx 1$, practically no information is involved, while if p_i is very small, very much information is involved. Using the logarithm gives us also a useful property of information being additive. Although any base could be used, the logarithm to the base two is normally used, particularly where binary coding schemes are under consideration. The information is then said to be measured in bits (binary digits).

It follows, then, that the average information contained in selecting one message is

$$H = \sum_{i=1}^{N} p_i I_i = \sum p_i \log_2 \frac{1}{p_i} = -\sum p_i \log_2 p_i \quad \text{bits/message}$$

This choice of a measure of information content has been put on a much firmer mathematical basis by several authors. The above equation gives the entropy, H, of a set of probablities. It can be shown that H is a maximum when all the probabilities, p_i, are equal, and $H = \log N$. A more detailed development can be given to show that the concept of entropy can be extended to sequences of statistically correlated messages and to continuous message signals. However, for our purposes, we are normally interested in the information content of a finite set of messages, and furthermore, if we consider successive messages to be statistically independent, the above equation can be used directly to estimate the theoretical information content of a set of messages, once their relative probabilities are known.

Channel Capacity

The capacity of a binary transmission channel can easily be estimated. In this case, only two language characters, "Zero" and "One" are used. They are presumably transmitted as pulses of equal duration. The capacity of the channel can readily be defined as

$$C = \lim_{T \to \infty} \frac{\log n(T)}{T}$$

where n(T) is the number of distinct sequences of symbols of length T which can be expressed in the language of the channel. In this case, both symbols are of equal duration, t_o. In time T there are 2^k sequences, where k is the largest integer in T/t_o.

Then

$$C = \lim_{T \to \infty} \frac{\log 2^{T/t_o}}{T} = \frac{\log 2}{t_o} = 1/t_o \text{ bits per second}$$

Shannon's Fundamental Theorem for a Noiseless Channel

Theorem: Let a source have entropy H (bits/symbol) and a channel have capacity C (bits/sec.). Then it is possible to encode the output of the signal source in such a way as to transmit at the average rate $C/H - \epsilon$ symbols per second over the channel, where ϵ is arbitrarily small.

In order to achieve the theoretical maximum transmission rate in a practical case, rather complicated codes would usually be required. Nevertheless, the theorem provides us with a standard of comparison by which to judge the amount of redundancy in a given coding system .

Minimum Redundancy Codes

Before stating some general rules for the construction of simple codes, the following examples are given:

Example 1:

Suppose that the following set of messages is to be coded

Message M_i	Probability P_i	Standard 2-bit Code	Minimum Redundancy Code
A	0.55	00	1
B	0.25	01	0 1

(cont.)

Cont. Message, M_1	Probability, P_i	Standard 2-bit Code	Min. Redundancy Code
C	0.15	10	0 0 0
D	0.05	11	0 0 1

The average entropy of the message set is

$$H_{av} = -\sum_{i=1}^{4} p_i \log p_i = 1.62 \text{ bits/message}$$

A standard 2-bit code would require 2 bits/message. Note, however, the minimum redundancy code. This code is so constructed that the message can be transmitted in any order without ambiguity, and without requiring a mark or space separating individual messages. The average message length is

$$L_{av} = \sum_{1}^{4} p_i \, l_i = 1.65 \text{ bits/message}$$

This represents a 17.5 percent reduction in average message length, when compared to the standard two-bit code, and is close to the theoretical lower limit. Consider another case:

Let the set of messages to be coded be the following:

Message, M_i	Probability, p_i	Standard 2-bit Code	Min. Redundancy Code
A	0.70	01	0
B	0.20	10	10
C	0.10	11	11

Here the average message entropy is 1.15 bits, and the minimum redundancy code for individual messages has an average length of 1.3 bits. However, by coding in message-pairs the following result is achieved:

Message-Pair	Probability	Code
AA	0.49	1
AB	0.14	000
BA	0.14	001
AC	0.07	0100
CA	0.07	0101
BB	0.04	0111
BC	0.02	01101
CB	0.02	011000
CC	0.01	011001

And the average pair length is 2.335 bits, which is not too far from the theoretical minimum of 2.30. The process can, of course, be extended by forming longer groups to achieve more efficient coding. Once a given set of messages, or message groups has been established, the construction of a minimum redundancy code can be done in a straight-forward manner, as outlined in the following section.

Construction of Minimum Redundancy Codes*

A brief procedure for forming minimum-redundancy codes is given below.[38]

1. Arrange all possible messages in order of decreasing probability. Assign as the last digit in the coded output a 0 to the next-to-last message, and 1 to the last message. These two messages will subsequently agree in all the (as yet unknown) digits preceding the last. They are some-

*Sometimes called Huffman Codes. See Reference 43.

times said to agree in their prefix.

2. Merge the last two messages to form a new message, adding their probabilities. Repeat Step A. Continue until all messages have been merged.

A characteristic of this type of coding is the prefix property; i.e. no codeword is a prefix of any other longer codeword. This property guarantees that any sequence of codewords can be written down in any order without spacing and still be uniquely decoded into a sequence of source symbols.

Other forms of coding, notably the Shannon-Fano system, yield straightforward methods for reducing the redundancy of a coding system, however, the above method will always produce a minimum-redundancy code, whereas others may in some cases fail to do so. The above method was used to produce the codes given in the above examples.

Control of Source Rate

Note that above technique implies some means of controlling the source rate, so that the generation of codewords can proceed at a constant bit rate. In many cases the short-term information rate of a given sequence of messages may be higher than average, and some form of buffer storage of messages is required, so that the messages can be coded and transmitted at a lower average rate. Statistical techniques can be applied to estimate the amount of storage required to insure adequate interim storage of the messages prior to coding, where the message source rate cannot be controlled.

Coding of Rate and Random Events ($H < 1$)[24]

An important category of messages not yet considered is one where one message in the set is much more probable than any of the others. In general the average entropy H, is less than one. For this case, the coding methods discussed above are not useful, since each message terms requires at least one code digit, giving $H > 1$. If the message terms are statistically independent, then message sequences will consist of long runs of a single, high-probability symbol, interrupted by occasional low-probability symbols. Briefly, the approach to coding is to describe each run of high-probability symbols by a binary number that indicates the number of symbols in the run. If the number of bits in the run-length word is less than the run length itself, and if the other message terms are sufficiently rare, the total number of bits required to code a sequence of messages is reduced.

Predictive Coding[24]

This is a general coding technique wherein both the transmitter and receiver store information about past message terms, and from them, estimate the next message term. The transmitter sends not the message term, but the difference between it and the predicted value. This is a method which has usefulness in coding correlated message terms, without utilizing excessive amounts of transmitter and receiver storage. The technique requires consideration of higher-order conditional probability distributions of the messages and is considerably more complicated than any discussed in this paper. The value of such coding techniques is dependent upon the nature of the message ensemble to be coded.*

Error Detecting and Correcting Codes

The elementary coding methods discussed above suggest methods by which the number of bits required to specify a sequence of messages may be reduced. Since this reduces the output bandwidth of the communication

* In general, the technique can be applied to the bandwidth and energy limited ergodic processes. For a detailed discussion of predictor criterion, the reader is referred to the referenced article by Elias.

channel, the output signal-to-noise ratio is correspondingly improved. Also for a given transmitter power the bits-without-error capability improves. Note, however, that by removing a signal redundancy, the susceptibility of the remaining signal to a degredation by noise is actually increased.

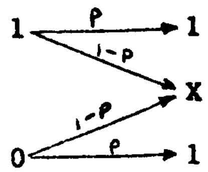
In comparison to redundancy reduction coding, error detection and correction coding methods add controlled amounts of redundancy into the code to allow accurate decoding of received messages containing missing or erroneous bits. The amount of redundancy required depends upon the desired accuracy and the average bits-without-error capability of the system.

The general theory of error detecting and correcting codes has been studied in considerable detail. Most of the work has been in developing methods for use with systematic codes, i.e. codes where each codword has the same length. The problem of incorporating efficient error-correcting redundancies of minimum redundancy codes of variable word length is somewhat more difficult. This limitation can be overcome to an extent by adding the extra bits to blocks of code words of equal length, a technique which again requires temporary storage of information at the transmitter and receiver. The following discussion assumes that the code words or blocks of codewords are of equal length.[11,41]

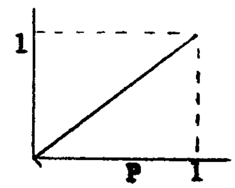
<u>Types of Errors</u>

1. Binary Erasure Channel. One form of error in a binary communication channel is one where received bits are entirely missing, as illustrated on the following page.

Binary Erasure Channel



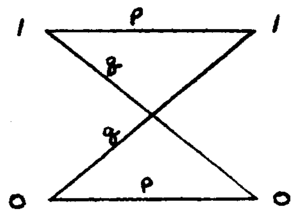
Information Rate



This is called a binary erasure channel. In the figure, it is assumed that input "1's" and "0's" are equally likely, and that they are erased with probability p. In this case, the average information received is p bits per input symbol.

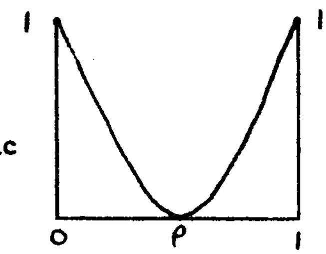
2. Binary Symmetric Channel. In the figure below, the channel also accepts only the input symbols 1 and 0, but it also only produces the same two output symbols. With probability p its output reproduces its input, also with probability $q = 1-p$, the output symbol is incorrect. In this case, the average information rate, R_1 equals $(1 + p \log_2 p + q \log_2 q)$ bits per symbol.

Input



Output

Symmetric Binary Channel



<u>Error-free Coding in a Binary Erasure Channel-Parity Check Coding</u>.

In a binary erasure channel, a simple parity check will detect and correct one erasure in a code word. A parity check is constructed by adding one digit to each code word. The added digit is selected so as to make the total number of 1's in a code word either always odd or even. Thus a single erasure can be corrected by the receiver by totaling the number of 1's in the received word. An additional check word may be added after each group of code words; this word would be constructed so as to form a parity check of the preceeding words by columns. By repeating the process of parity checking larger and larger groups of words, the error probability can be made as small as necessary.[38]

Error Correction of Binary Symmetric Channel

Note that single parity checks are of no value for this case, since there is now way of telling which bit in a word is inverted. A double parity check could locate a single error in a block of words, but in this case, parity checks are only able to detect errors, but not to correct them. A more general coding method is required, often called the Hamming code.* The general technique is summarized below:

1. To each group of message digits, k checking digits are added.

2. The k checking bits are added in such a way that when parity checks are made on the proper groups of received digits, not only will an error be detected, but the position of the error will be known, and designated by a position number formed during the parity checking.

3. The position number locates the error in any one of the (m + k) codeword positions. The required value of k is given by the following: $2^k \geq m + k + 1$. Normally, positions 1, 2, 4, 8, etc. are used for checking digits.

Extension of Error Correction and Detection Coding

The above checking process can be shown to have a geometrical interpretation. If we consider the possible combinations of $n = m + k$ digits as forming points in an n-dimensional space, then we define the "distance" between two points as the number of coordinate (or digit positions) by which two code words differ. In terms of error detection and correction "distance" implies:

* See reference 11 and 41

Minimum Distance	Meaning
1	Error not detectable
2	1 Error detection (Simple parity check)
3	Single error correction (Hamming code discussed above)
4	1 Error correction and 2 Error detection
5	Double error correction
etc.	

Again, the error correcting capability of the coding can be made as reliable as necessary at the expense of increased code word length. However, for most systems, a simple Hamming code (distance 3) is adequate. The number of check bits, k, required to check message bits is given by the inequality above, and is tabulated below:

K check bits	M message bits
3	$\leq$ 4
4	$\leq$ 11
5	$\leq$ 26
6	$\leq$ 57

Summary

From the above discussion, it may be seen that there are many factors to be considered in constructing codes for a given communication system. In the case of the OAO program, the data to be transmitted falls into two general categories:

1. Data about conditions in the satellite, for use in positioning and adjusting the instrumentation and for monitoring various command channels.

2. Primary data from the satellite instrumentation itself. This would, of course, include the signals generated by these TV cameras and spectrometers.

The data in the first category would presumably consist of a large number of slowly-varying quantities, and it could probably be transmitted best by conventional time multiplex PCM-FM techniques. A straight binary code word from an A/D converter for each data sample possible with a parity bit would be suitable. Transmission errors would not normally be serious, since the slowly changing nature of the data would allow interpolation of successive samples at the receiver.

On the other hand, coding methods for the primary data from the satellite instruments must be examined more carefully. Since there will be a large amoung of statistically varying data, and since good transmission reliability should be sought, somewhat more sophisticated coding techniques may be required to achieve reliable transmission over a narrow bandwidth with low transmitter power.

For each class of messages to be coded, the requirements may differ. In some cases, it may be desirable to remove a large amount of redundancy in order to reduce the required channel bandwidth. In other cases, controlled redundancy, in the form of parity check bits, etc. may be included in the construction of the code, in order to reduce the transmission errors. In the general case, both techniques may be applied simultaneously.

Obviously, coding methods should be kept simple, wherever possible, in order to reduce the complexity of the coding equipment required in the satellite. The advantages of bandwidth reduction coding schemes must be weighted against the overall goals of system reliability and simplicity.

VI. DETECTION AND DECODING

The selection of a detection and decoding system is generally made on the basis of a system "threshold." This "threshold" is usually defined as the point at which the system no longer performs within desired specifications. In telemetry systems the threshold is the minimum signal-to-noise ratio at which the system performs with the desired reliability.

With this criterion in mind the recently much published threshold improvement detectors were investigated. In particular the synchronous detectors using a phase-lock loop were considered since they apparently offer considerable improvement over conventional FM detectors.

When the discriminator alone is considered the phase-lock loop does have an advantage over the conventional discriminator at low signal-to noise ratios and thus improves the discriminator threshold. However, in a detection system each component may have its own threshold with one component limiting the total system threshold. In and FM/FM system the limiting component can be the discriminator and by improving its threshold the system threshold may be improved.

In the telemetry system under consideration here the data dictates the use of a PCM system, and in a PCM system it is the decoder and not the detector which is the weak link in the threshold chain. The circuit in the decoder which "decides" whether a given pulse is a "zero" or a "one" requires a minimum signal-to-noise ratio at its input to make this decision with a given accuracy.

McRae[58] shows that the curve of "bits without error" when plotted versus signal-to-noise input to the decision circuit is asymptotic at about 16 DB. Therefore, in PCM little is gained by reducing any threshold below

16 DB unless the decision circuit threshold is also lowered. At signal-to-noise ratios in this region conventional discriminators have advantages over phase-lock loops because of their reliability and simplicity.

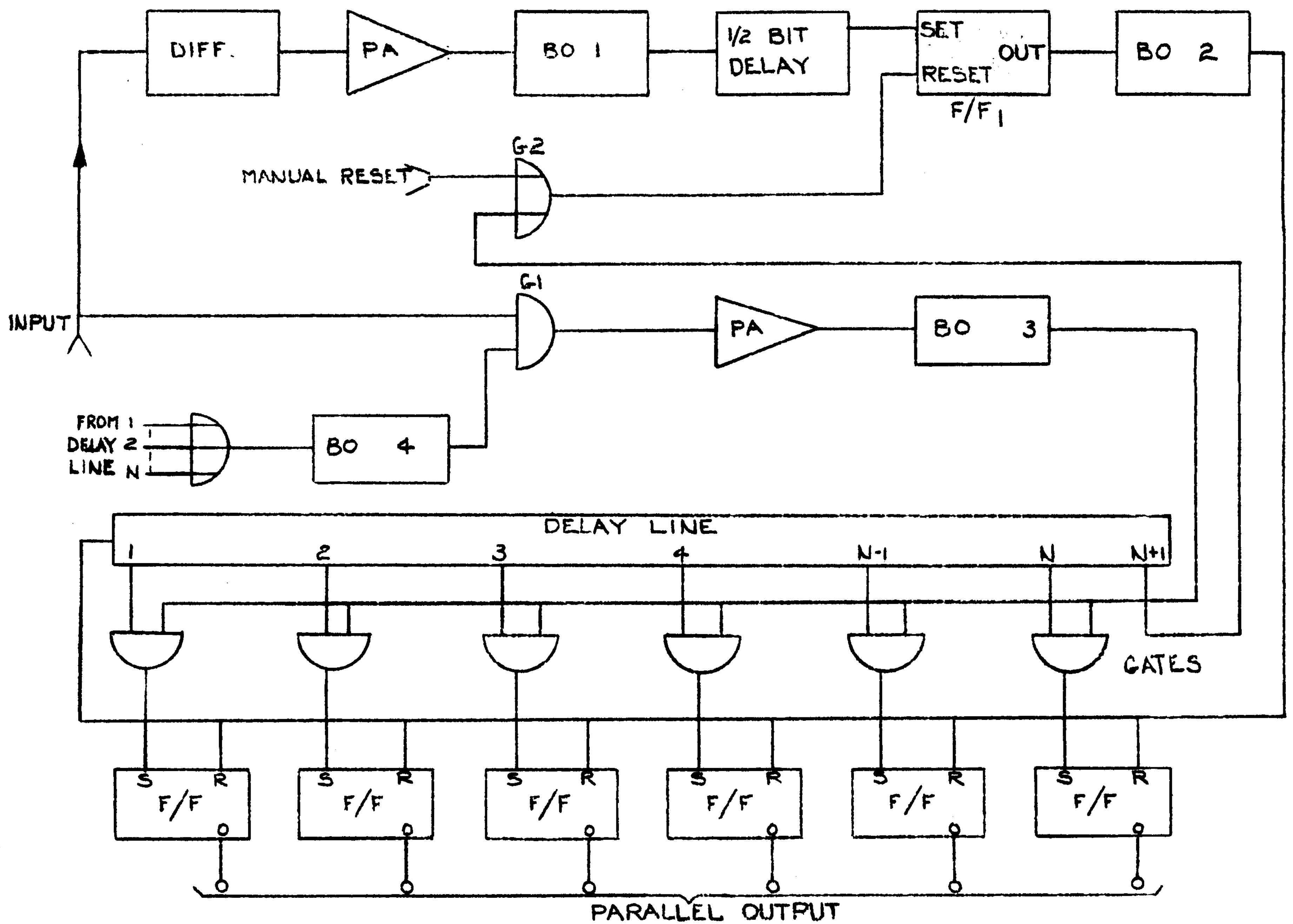
In PCM systems the decoder is a series to parallel converter. The function of the series to parallel converter is to separate the serial train of pulses out of the receiver into words and present them in parallel form to the recorder and quick look analog decoder. To accomplish this the converter must make the decision as to whether each input pulse is a "one" or a "zero". This requirement makes the converter by far the most important single component in the entire satellite system. Nothing can be done after this decision is made to improve the accuracy of the data transmission. Therefore the selection of this unit should be made only after very careful consideration.

In addition to the "one" or a "zero" decision the converter must be able to recognize the beginning and end of each data word and each data frame. This is necessary to insure that each data word contains data sampled from only one channel and that the channel from which the data came is known.

There are several schemes for establishing and maintaining this synchronization. Their complexity is in general a function of the input signal-to-noise ratio and changes in data bit rate. For a system with fixed bit rate and reasonable signal-to-noise ratio (15 DB) the converter can be made extremely simple and reliable. One of the more straight forward schemes is described by Seaver[78]. His description is essentially repeated in the paragraphs that follow.

To provide a code for the synchronization two extra bits are added to each word and an extra word provided at the beginning of each frame. Figure XVI is a block diagram of Seaver's[78] series-to-parallel converter. This system uses a "zero-to-one" transition for word synchronization, that is, each word starts with a "one" in its first bit and ends with a "zero". Thus as a word ends and a new word begins there is a "zero to one" transition.

In the diagram the input pulses enter G1 and the differentiator; G1 is an "and" gate and will not pass any pulses unless there is also a pulse out of B04. When a "zero to one" transition occurs at the input, the output of the differentiator is a positive pulse which triggers B01; this in turn sets F/F1. When F/F1 is set, B02 triggers and resets all the flip flops in the parallel read out; the pulse from B02 also enters the delay line. The time spacing of the taps on the delay line is the same as the time spacing of the data bits. Now as the next data bit enters the system B04 is triggered since the pulse from B02 appears at the tap No. 1 at this time. The pulse from B04 and the input pulse are coincident at G1 and it opens, triggering B03 which sets the first flip flop on the parallel output. B04 continues to trigger each time the pulse in the delay line passes a tap and each time B04 triggers, B03 is triggered if a "one" appears coincident at G1 with the pulse out of B04. This process continues for the "N" pulses which make up the bit word. When the pulse in the delay line passes the (N + 1) tap it resets F/F1 through "or" gate G2. If the "zero to one" transition which started the decoder was at the beginning of a word, as it should have been, there will be another "zero to one" transition present at the input to the differentiator at this time, and F/F1 will again be set and trigger B02. When B02 triggers it will reset all the output flip flops and start down the delay line to decode the next word entering the decoder. As B02 resets the



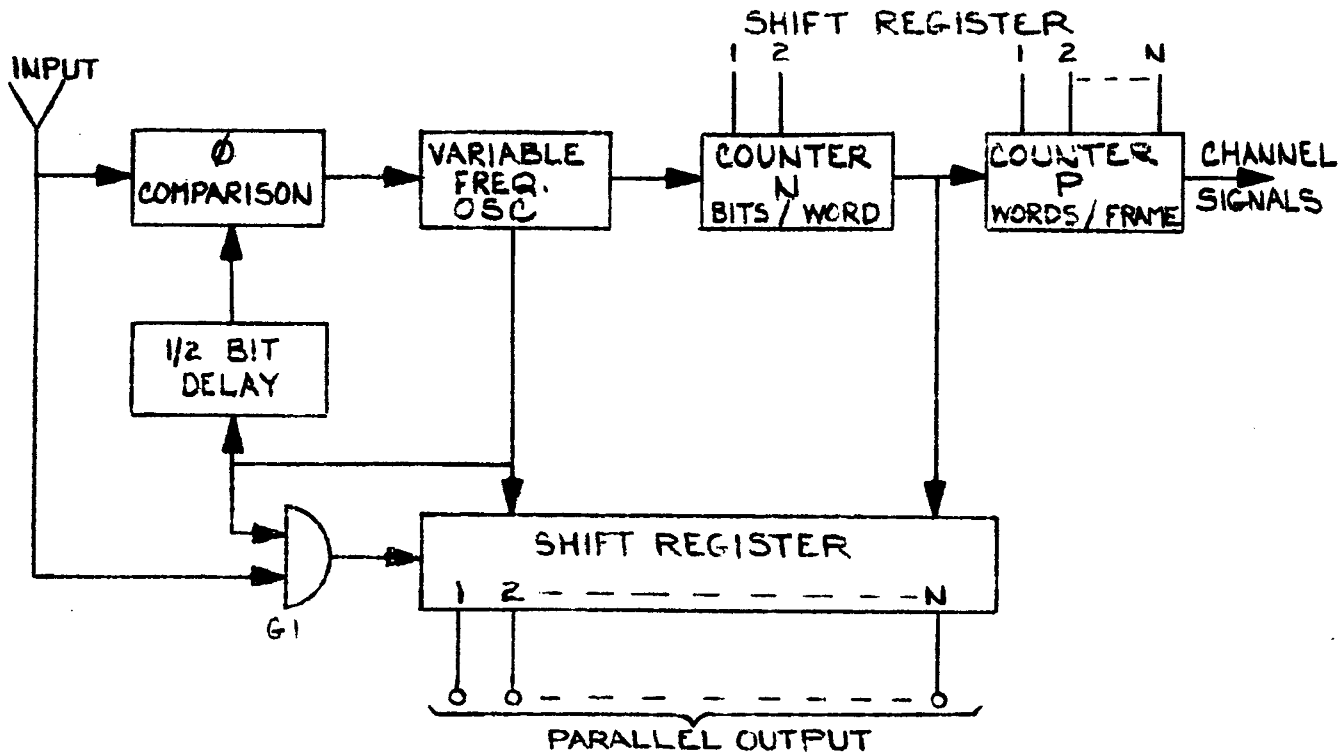
Delay Line Series to Parallel Converter
Figure XVI.

output flip-flops, the parallel word is read out of the flip-flops, completing the serial to parallel code conversion. If a combination of data should appear such as to cause a zero-to-one transition to occur at the same point in a number of successive words, it would be possible for the decoder to lock on to this series of transitions, rather than the desired word sync, giving a scrambled word composed of parts of two input words. This problem can be avoided by using a coded word to identify the beginning of a frame which has only one "zero", and that in the last bit in the word. Now if the decoder locks on to a "zero-to-one" transition in the middle of a word the sync can continue only until the code word at the beginning of the frame arrives at the input. This word will not have the "zero-to-one" transition except at the end of the word and the differentiator will wait for this transition, thus locking the decoder on to the beginning of the first word in the new frame. Only the data in the first frame is lost; this will usually happen, since this type system obtains its sync on frame beginning.

This type system has several very real advantages; probably the greatest is its simplicity, which usually goes hand in hand with reliability. The chief disadvantage of this type system is the differentiator on the input; this requires a fairly good signal-to-noise ratio for reliable operation. This is a very important consideration in any system receiving transmission from a low power satellite system.

An alternate synchronizer system is one which uses a shift register for the parallel conversion, and an oscillator and counters to maintain synchronization.[78]. Figure XVII is a rough block diagram of such a system.

In this system the pulses out of the oscillator are delayed 1/2 bit, and their phase is then compared with the input pulses. The phase comparison



Shift Register Series to Parallel Converter

Figure XVII.

circuit varies the frequency of the oscillator to maintain phase synchronization with input pulses. The output pulses from the oscillator also shift the register and act as strobe pulses to admit "ones" to the register through gate G1.

The word and frame synchronization is obtained by two counters, one counter-to-(N) for the (N) bits in each word, and one counter-to-(P) for the (P) words in the frame.

The word counter looks at all (N) entries in the shift register, waiting for the particular unique code word which identifies the start of a new frame. Once the code word is identified the counter starts to count, and the circuit which recognizes the code word is disabled until the counter counts to (P). When the counter has made (P) counts, the code recognition circuit is again enabled and looks at the word present in the shift register. The code word should be present in the register: if it is, counting continues: if it is not, the identifying circuit continues to examine each successive word entering the register in its search for the code word. During this searching interval none of the words which are dumped out of the shift register are routed to any data channels by the word counter. When the code word is again found normal counting begins.

The counter-to-(N) which counts the (N) bits in each word monitors two positions in the register seeking the word identification code. When the code is found, the identification circuit is disabled while the counter makes (N) counts. After (N) counts the identification circuit is again enabled and seeks the code which should be present at this time. At the end of each (N) counts the bit counter gives one count to the word counter and dumps the shift register. If the identification circuit in the word counter has enabled the counter (because it has found the code word indicating the beginning of the frame), it will accept the count from the bit counter and will also route the output of the shift register to a data channel.

When the data is being recorded for later processing, it may be possible to make use of the data contained in the first partial frame. However, in command system data it is imperative that no data be routed to any channel until the word counter has made frame synchronization. The result of the wrong command getting into a command channel could be disastrous.

The important element in either of the two systems described here is the element which actually decides whether or not a given input pulse is a "zero" or a "one". In both Figure XVI and Figure XVII, it is G1 which makes the decision.

Often these systems are used with pulse shaping circuits on the input; when this is done all the advatage of the strobing is lost since the "one" or "zero" decision is made in the shaping circuit.

The decision circuit G1 must have a 16 DB signal-to-noise ratio for reliable operation in either system, and the differentiator operation should be very reliable with this signal-to-noise ratio. Therefore, the system in Figure XVI appears to have considerable advantage over the one in Figure XVII, when the bit rate is fixed.

VII COMMAND SYSTEM

In considering the command system, the first and most obvious requirement is for absolute reliability. Since the ground station transmitter can presumably supply whatever power may be necessary for error-free transmission, such considerations as bandwidth and power conservation become strictly secondary to the overriding requirement for reliability. On this assumption, that the command signal can always be delivered at an adequate signal-to-noise ratio to insure error-free reception, the primary factors to be considered in a design seem to be as follows: (1) to insure the minimum likelihood of incorrect interpretation of the signal, (2) to insure that an incorrect interpretation of a command will not result in any permanent impairment of the satellite performance: (3) to insure that the failure of any given component will have the least possible affect on the performance of the satellite.

It is obvious that the above requirements imply keeping the command system as simple as possible. Unfortunately, there are so many different functions to be controlled in a typical OAO satellite, that the command system required will of necessity involve a degree of complexity exceeding that of any satellite yet launched. The first problem is the matter of access to the various control channels. The problem basically resolves to a choice between random or sequential access to each channel. When the number of channels is small, random access can be simply achieved by tone signalling, assigning a separate tone to each channel. However, in the case of the OAO satellites, there may be 60 or more separate control channels, so that random access would require some form of digital addressing. This in turn will involve some type of serial to parallel converter and a decoding matrix, both relatively complex devices. Since the entire operation of the satellite depends on the commands getting to the right place, this seems to violate requirement (3) above. However, sequential access is simple to implement by routing the command signals to the various channels through gates under the sequential control of a shift register activated from the ground station. The failure of a gate would affect only the channel in which it was located, and magnetic shift registers can be made extrememly reliable, thus making the likelihood of a single failure affecting the entire satellite quite remote. As for speed of access, even if it were necessary to step through nearly all the channels to get to the desired one, the access time would still require only a fraction of a second, using a bit rate in the kilocycle range. Thus, random access seems to have no advantages whatever, and sequential access is therefore recommended.

Now, assuming that the signals can be routed to the proper channel, the next problem is what type of signalling should be used to convey the commands themselves. Some of the commands will certainly be little more than ON-OFF, for which tone signalling would be adequate. However, for some channels, particularly the telescope controls, many levels of control may be required, so that digital signalling seems required. Assuming a high signal-to-noise ratio, as indicated above, the largest source of error in digital signalling is improper synchronization. However, synchronization is only a problem with non-return-to-zero coding, whereas return-to-zero coding is basically self-synchronizing. The only serious objection to return-to-zero coding is that it doubles the bandwidth requirement for a given data rate, which is not a serious drawback in this case.

As to the actual codes to be used, the high signal-to-noise ratio available together with the inherent reliability of return-to-zero coding obviate the necessity of error-checking or correcting codes, so straight binary coding is recommended. The word length required will depend on the precise nature of the devices to be controlled.

For the actual transmission of the commands, three key signal levels are required for the return-to-zero command signals, plus a fourth level to pulse the address shift register. This register could be controlled by some special combination of command codes, but it seems simpler and more direct, and therefore more reliable,to use a fourth level. This four level transmission could be accomplished by practically any of the conventional modulation techniques, but two methods which suggest themselves as particularly simple, reliable, and highly developed, are frequency-shift keying and tone-keyed FM.

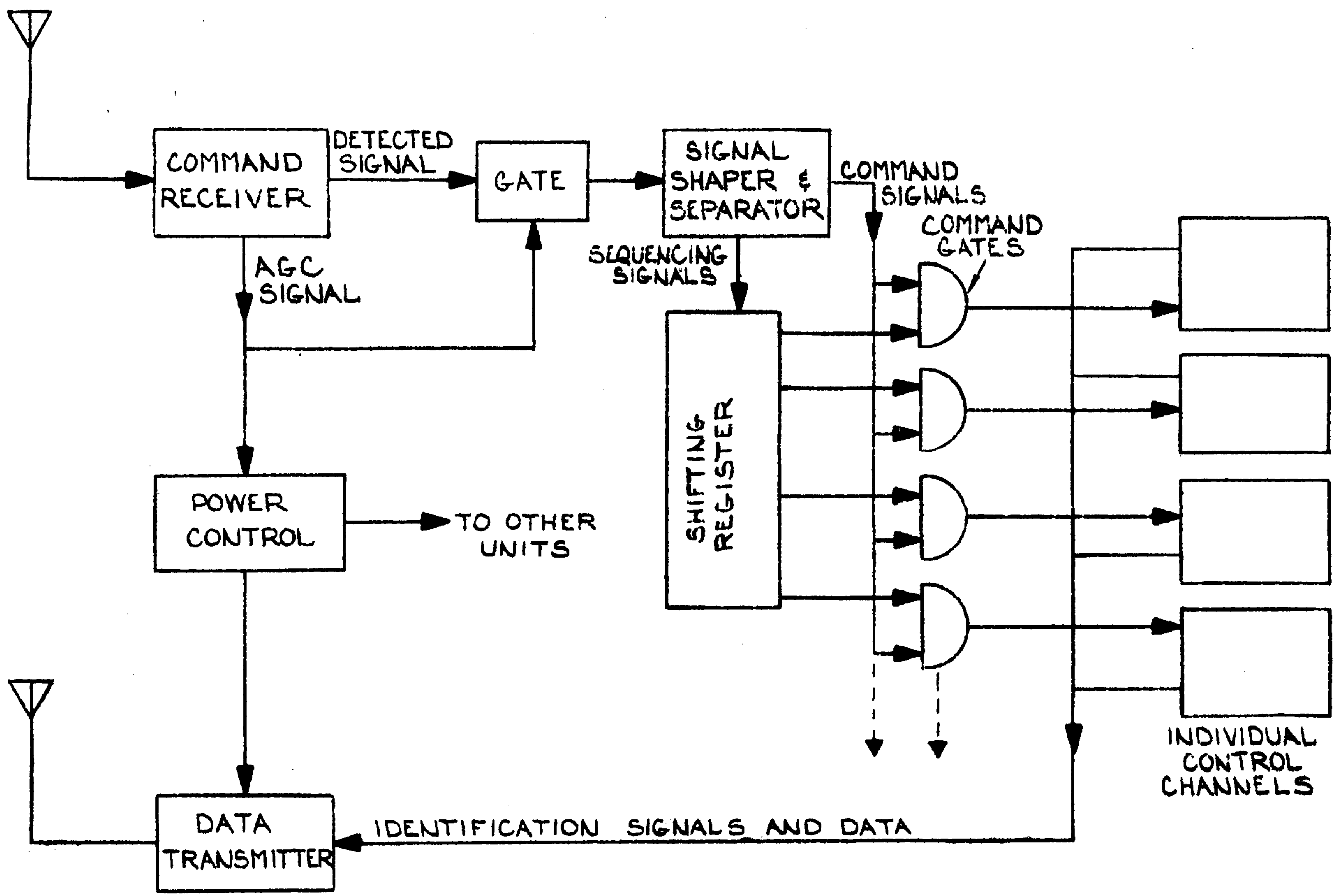
Although it is not a major problem here, it is certainly of interest to examine the power and sensitivity requirements of the command link. As mentioned elsewhere in this report, the power limitations for a transistorized output stage on the satellite transmitter will limit the bit rate to about 10 kc. Although this limitation does not apply to the ground transmitter, it will be assumed for the sake of clock system simplicity that the same bit rate will be used throughout the satellite system. Using RZ coding then, the RF bandwidth required is 20 kc. The carrier frequency will be assumed to be in the vicinity of 150 mc. With plus-minus transmission, the required signal-to-noise ratio at the receiver is generally taken to be from 12 to 15 DB. In this case a high signal-to-noise ratio is desired to insure error free transmission, so the receiver signal-to-noise ratio will be taken as 24 DB. Other system values are, receiver noise figure 6DB; receiving antenna, isotropic; fading allowance, 20 DB; transmitting antenna, 10 ft parabolic; path distance, 2000 miles. With these values, the power calculations work out as shown below.

Receiver noise Level	-125 DBM
Req'd signal to noise ratio	24 DB
Receiver sensitivity	101 DBM
Path loss	147 DB
	46 DBM
Fading allowance	20 DBM
Radiated power req'd	66 DBM
Transmitter ant. gain	11 DBM
Transmitter power	55 DBM = 330 watts

These requirements can obviously be met with no difficulty whatever. In view of the desirability of keeping the satellite equipment as simple as possible. It would probably be desirable to increase the transmitter power considerably, thus reducing the requirements on the satellite receiver. For the transmission from the satellite, for which the regular data transmitter would be used, the computations will be the same, except that a larger dish will be used for the receiving antenna at the ground station. If a 60 ft parabolic antenna were used, the power requirement for the satellite transmitter would be approximately 4 watts.

Satellite Section

The overall arrangement of the satellite section of the command system is shown in Figure XVIII. The normal factors considered in establishing the diagram and the sequence of operation are described below.



Block Diagram - Satellite Command System

Figure XVIII.

When the satellite is out of communication range, it is desirable to insure that noise signals cannot cause a response in the command channel. In addition, various devices which have a function only when the satellite is in communication range should be shut off to conserve power. Both functions can be accomplished through the use of the AGC signal of the command receiver, as shown in the figure. When the satellite is out of range, the absence of AGC will cause the gate at the output of the receiver to close, as well as turning off the power to the data transmitter and various other units. (It is assumed here that the beacon transmitter for the tracking system is a separate unit that will operate continuously unless turned off by a signal from the ground).

When the satellite comes into range and the ground station starts transmitting, the presence of the carrier signal will be indicated by the AGC signal, resulting in the receiver gate opening and the transmitter being turned on. When the signal from the satellite transmitter is received at the ground station, this will provide an automatic indication that the satellite is ready to receive commands. Note an additional feature here, that by requiring a certain minimum AGC level, it is possible to insure that the satellite will respond only when the signal is strong enough to insure error-free reception.

When the received signal from the satellite indicates that it is ready, the ground station can then proceed with the desired operations in sequence. Each pulse to the shifting register will shift command to the succeeding channel which can then be interrogated or commanded as desired. Since there will be so many different types of devices to be controlled, there will be no attempt here to consider the individual control channels except to note that each channel should be provided with means for sending an identification signal on command, to verify that the commands are going to the correct channel.

In summary, it will be seen that there is nothing particularly sophisticated or complicated about the satellite command system. This is in agreement with the objectives discussed earlier. Every attempt has been made to keep the system just as simple and straight forward as possible, consistent with the requirements. Many more sophisticated and elegant techniques were considered and rejected as being neither necessary for nor consistent with the design objectives.

VIII. GROUND STATION

General Organization

In this section the organization of the entire ground station will be considered, not only those parts which are directly part of the command system, but also those parts which are only indirectly connected with the command system. This is necessary because the various parts of the ground station overlap too much in their functions to be separated into independent pieces.

The ground station must provide facilities for quick, flexible, and reliable control of the satellite, plus facilities for recording and display of the data received from the satellite. Although the interpretation of the data is a separate function, and would presumably be handled at a separate computing center, the arranging and grouping of the data into a form compatible for computer entry will be considered a function of the ground station. From a standpoint of speed and reliability, it is desirable that these operations be carried out under the direction of a computer following a stored program. In addition, there must be provisions for the operator to take manual control in the event of malfunctions or other difficulties.

The general sequence of operations will be as follows. While the satellite is out of range, the operator will store in the computer memory the command program to be carried out. This program will consist of instructions as to which channels are to be controlled, the exact commands to be sent, data to be recorded or displayed, etc. When the tracking system indicates that the satellite has come into communication range, the operator will start the data transmitter. The return signal from the satellite will indicate that the satellite is ready to receive commands, and will automatically start the computer on its control program. Using the return signal from the satellite will provide a safety feature in stopping the command cycle if communications with the satellite are interrupted for any reason. It should be noted that using the receiver AGC signal as a control both on the satellite and at the ground station requires that both units must be transmitting and receiving simultaneously at all times. This will require more complexity in the communications system than using a "send or receive" system, but the extra reliability obtained by maintaining continuous contact in both directions would seem to justify this extra complexity.

The computer will now proceed with the control program, reading out the data, setting up new operations, reorienting the satellite, etc. Under normal conditions the computer should be able to carry out the complete program without human interference, and when the program is complete, will turn off the transmitter, and return to a standby state, awaiting a new program.

In the event of trouble, the computer should sound an alarm of some sort, and stop operations to wait for corrective action to be taken by the operator. The operator must be able to interrogate and command any given channel directly. He then may take corrective action as necessary, and return control to the computer. In some cases, the difficulty might be such as to render the stored program ineffective, in which event the operator could complete the program manually or start over, or take whatever action seemed appropriate. In some cases, such as identifying a star field, it might be necessary for the computer to stop and wait for a decision from the operator before proceeding.

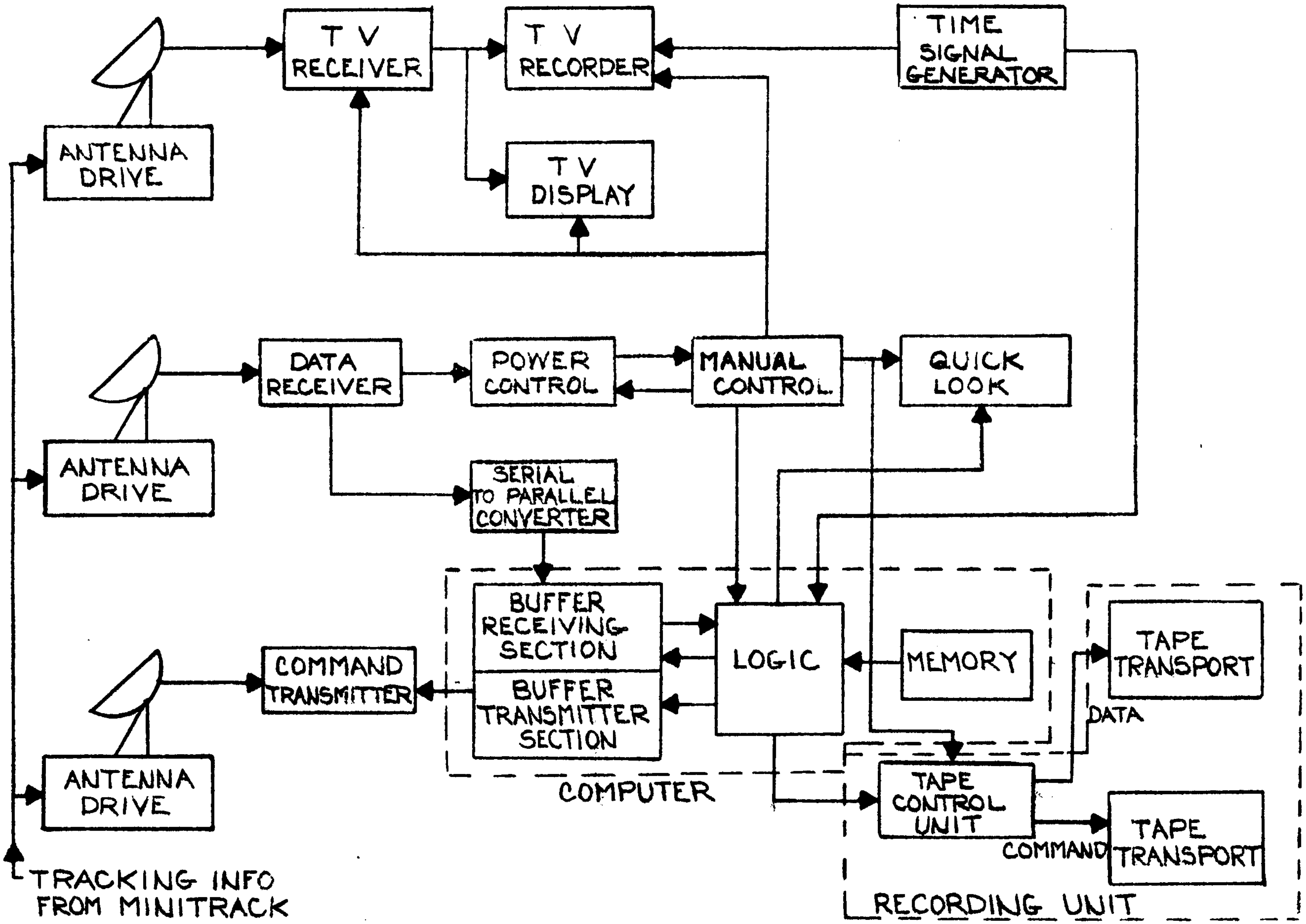
The overall block diagram of a ground station which would be capable of all these functions is shown in Figure XIX. The various individual blocks will be discussed below.

<u>Computer Section</u>

The computer will basically consist of a storage, or memory, section, in which the program will be stored, a logic section which will compare the actual state of a channel with the desired state, and originate the necessary commands, and a buffer section which will translate the computer signals to a form suitable for use on the satellite, and vice versa. A typical sequence of operation is described in the following paragraphs.

When the signal from the power control section indicates that the satellite is ready to receive commands, the number of the first channel to be controlled will be transferred from storage to a register in the logic, and the buffer will be instructed to send out an interrogation signal to determine which channel is under control. When the number of the active channel is received, the buffer converts it to computer form and sends it to the logic, where it is compared to the desired channel number. If the active channel is not the desired channel the logic determines by subtraction how many pulses will be needed to sequence the shifting register on the satellite to the desired channel. The logic then directs the buffer to send out the proper number of sequencing pulses, and when they have been sent, calls for another interrogation to verify that the correct channel is now active.

When the right channel is gated, the logic then calls up the command from the memory. In the normal situation, the first step will probably



Block Diagram - Ground Station System

Figure XIX.

be determining the present state of all control channels and instruments. If this is the case, the command will be a data interrogate signal, calling for the control device or instrument to report on its condition. Whenever this command is sent, the logic will also direct the receiving section of the buffer to route the data to the recording section, and the display section if desired. When the data is recorded, the buffer will so notify the logic, which will then move on to the next channel and repeat the sequence.

Another type of control situation will be the adjustment of telescope controls or the repositioning of the satellite. The exact sequence of commands will depend on the design of the various actuating elements, a factor not considered in this study. However, one method of attacking this problem would be to have the various controls driven by stepping motors which would move the control some incremental distance each time a pulse is received. In this case the procedure would be for the computer to select the proper channel as before, compare the present state with the desired state, compute the number of steps required, drive the device this number of steps, then compare actual position with desired, and repeat as necessary, until the desired position is achieved. If the desired position cannot be reached after some specified number of tries, the computer should stop and notify the operator.

<u>Manual Control</u>

Manual control, as indicated before, is required for two purposes, (1) to give the operator direct control over the satellite in the case of difficulties; (2) to provide for direct insertion of commands when the program requires the use of received information as a basis for decisions too complex for the computer to make. As an example of this second case, aiming of the telescope may be based on the recognition by an astronomer of a star-field in the finder telescope. (The Smithsonian group has had some success in implementing even this problem with a computer, but it seems likely that situations of this general nature are bound to arise in the utilization of these complex satellites) To accomplish these purposes about all that will be needed are a couple of buttons for sending sequencing pulses, to sequence the gating shift register or to step the control actuators, and a keyboard for inserting coded commands or addresses. In addition controls will be required to block out automatic action of the computer entirely and to take direct control of any devices usually controlled by the computer.

Display Section

The display section, which should, of course, be located in such a position as to be easily viewed by the operator while at the control console, should present two pieces of information -- the identification number of the channel under control, and the reading of the instrument or indicator device in that channel. For this purpose, two lines of decimal readout, of the Nixie type, would do the job very nicely. In addition, it would probably be desirable for trouble-shooting purposes to be able to display various signals in the computer.

In addition to the type of display already mentioned, there may be a requirement for various types of quick-look equipment. For example, the astronomers may want a running graph or tabulation of the photometric readings in a certain part of the sky. For this the computer must have the capability of selecting the desired data and routing it to the plotter or tabulator or other type or readout device.

As will be noted in the diagram, the television display and recording is entirely separate from the digital data channel. There are two reasons for this. First, although the television system has not yet been designed, it is quite possible that the signals used to transmit the television image will be entirely different from the digital signals used to transmit other types of data. Second it may be desirable to be able to view the television images continuously during command cycles. The television section of the ground station will be considered in more detail in the section on the television system.

Recording Section

The prime purpose of the recording system is of course to record the astronomical data sent by the satellite. The simplest approach would be to take the data directly from the receiver and put it on tape. However, as mentioned earlier, the data eventually has to be put in a form suitable for computer. In most telemetering systems the traditional approach has been to record the data directly as above, and when the data taking is complete, play the tape back into a device called a computer format generator. This device puts the data into the proper word lengths, inserts parity bits, and carries out whatever other operations are required, to produce a data tape suitable for computer entry.

This format generator is basically nothing more than a small computer itself, and since there is already a computer in this system, it seems only logical to let this computer do the tape preparation. Such an approach has two apparent advantages. First, it will require less apparatus. Second, it will save time by making in-line tape preparation possible, thus eliminating the playback cycle. All that is required is to insure that the computer has sufficient extra capacity over that necessary for the command operations. Accordingly, this is the approach that is recommended. The data signals coming from the serial to parallel converter (discussed elsewhere) will be sent to the buffer, which will adjust the level and shape of the pulses to suit the computer. Then the signals will be sent to the computer which will carry out the operations discussed above. The final output will be a tape ready for entry into whatever large scale computer is to be used for data analysis.

In certain circumstances it might be desirable to record the commands sent to the satellite, as well as the data received from it. The need for such recording is not as obvious as that for data recording, but in the case of trouble with the satellite, an exact record of all signals sent out might be invaluable in analyzing the difficulty. Since the computer is already controlling the tape recording, it will obviously be no problem to program the computer to place the commands on tape. Although it would be possible to place the command data on the same tape as the received data, it will be more convenient to have it on a separate tape. Since the cost of adding an extra tape transport should not be large, it is recommended that this be done.

An additional part of the tape preparation task is the insertion of time reference signals. For this purpose, a time signal generator will feed time signals on the tape at certain intervals as specified by the program.

Summary

The central element in the ground station is quite obviously the computer. The computer facilities required here are quite modest by modern standards. This entire ground station could quite easily be built around any one of a number of small general purpose computers now available, such as the Computer Data Corporation 160, the Bendix G-15, the Royal McBee LGP 30, the Burroughs E101, etc. The CDC 160, for example has a 4000 word magnetic core memory, several adding and subtracting registers, an arithmetic unit, decimal presentation of the contents of the registers, manual entry to the registers, and various other features. Program entry can be made by punched card, punched tape, electric typewriter or magnetic tape. Output can be to any of the above devices, and when equipped with an auxiliary tape control unit, the computer can carry out tape preparation in the manner discussed above. The other units listed above are similar, and thus the ground station could easily be built around any of them. It is recommended that this be done, rather than trying to design a complete new unit just for this purpose. The buffer unit will have to be built especially for the purpose since its exact characteristics depend on the computer and the communication system adopted. There are several companies, Telemeter Magnetics, for example, which manufacture buffer units for any conceivable situation. Thus the entire ground station portion of the command system can be built up in a straightforward fashion from commercially available apparatus.

Television

The television requirements for this project definitely break down into two distinct categories. First, there is a requirement for television to be used in gathering astronomical data, and second, a requirement for television to be used as an aid in aiming and guiding the satellites. The first requirement arises in the Smithsonian and Michigan units, the latter probably in all the units.

In the Smithsonian satellite, television will be used for gathering two distinct types of data. First, the experimenters wish to make a complete sky survey in certain ranges of ultraviolet by simply taking television pictures of sky with tubes sensitive only to the various ultraviolet ranges of interest. Since they are interested not merely in stars, but also in various types of extended sources, the complete television image must be transmitted. For this reason, essentially conventional television techniques must be used. To attempt to digitze the information in all 250,000 picture elements would result in a prohibitively large bandwidth. Even using conventional analog transmission, it will be necessary to use slow scan (one frame per second, or less) in order to keep the bandwidth down to a practical level. If the system otherwise exactly followed commercial standards, a one frame per second scan would require a 150 kc bandwidth for transmission. However, the tremendous ranges of light intensity involved (10^6 or greater) will necessitate special techniques which may lead to higher bandwidths. This problem is presently being studied by the Smithsonian group, and there is little more that can be said about it at this time.

The Smithsonian satellite also requires television for transmitting slitless spectra of certain stars. Since such spectra are essentially continuous images, this application will also require television of the type discussed above, except that the intensity ranges will not be as great.

On the Michigan satellite, the television system will be used as a spectroheliograph, i.e., it will be used to take pictures of various areas of the sun in certain wavelengths. Again the picture will be continuous, and television of the type discussed above will be required. There are problems in developing pickup tubes and optics for these specilized applications, but these problems are not the concern of this study. It is to be noted here that RCA has already developed and built a slow-scan TV for exactly this purpose, which has been successfully used with balloon telescopes.

For use in aiming the satellite telescopes, the television tube will view the field of a finder telescope which will cover a 15-degree area of the sky. The image transmitted to the earth need contain only enough information to enable an astronomer to recognize the star field. For this purpose it will be sufficient to transmit the position and brightness of all stars out to the sixth magnitude, to 1/2 magnitude accuracy. A study of star atlases indicates that the maximum number of stars of sixth magnitude or greater to be found in a 15-degree field is about 150, with the average number being 50 or less. Using a typical vidicon tube with 500-line resolution (250.000 picture elements) the size of a star image with telescope of the f/ numbers to be used here is slightly less than the size of a single picture element. Thus, there will be information in no more than 150 out of 250.000 picture elements, so that it is obviously inefficient, from an information point of view, to use a conventional TV system, transmitting information about all picture elements. Transmitting information only about those elements which contain information will obviously save a great deal of channel capacity, and it will certainly be desirable to do so if it can be done without a prohibitive increase in complexity in the transmitting equipment. A possible scheme for accomplishing this objective will be outlined in the following paragraphs.

The system to be described employs a stop and go scanning technique. The scanning beam of the camera tube will start scanning at a high speed in the normal manner, but there will be no output signal until a star image is encountered. When such an image is encountered, the scanning beam will stop and information on the position of the image and its intensity will be converted to digital form and transmitted to the ground station. As soon as the information has been transmitted the beam will resume scanning at a high speed, and the above process will be repeated at each star image. An advantage of this system is that the frame scan time is not fixed, so that it is not necessary to choose a bandwidth in accordance with the maximum information content of a frame, with the result that the system is operating well below capacity most of the time. Instead, the bandwidth will be set by the bit rate, which can be set as best suits other design factors, and the information content of a single frame will merely govern the length of time it takes to scan one frame. And, as will be shown below it does not require an especially fast bit rate to provide a scan rate quite acceptable for this application.

The block diagram of a possible system for implementing this technique is shown in Figure XX . Pulses from a master clock oscillating at 250 kc are fed to a binary scaler through a normally open gate. The

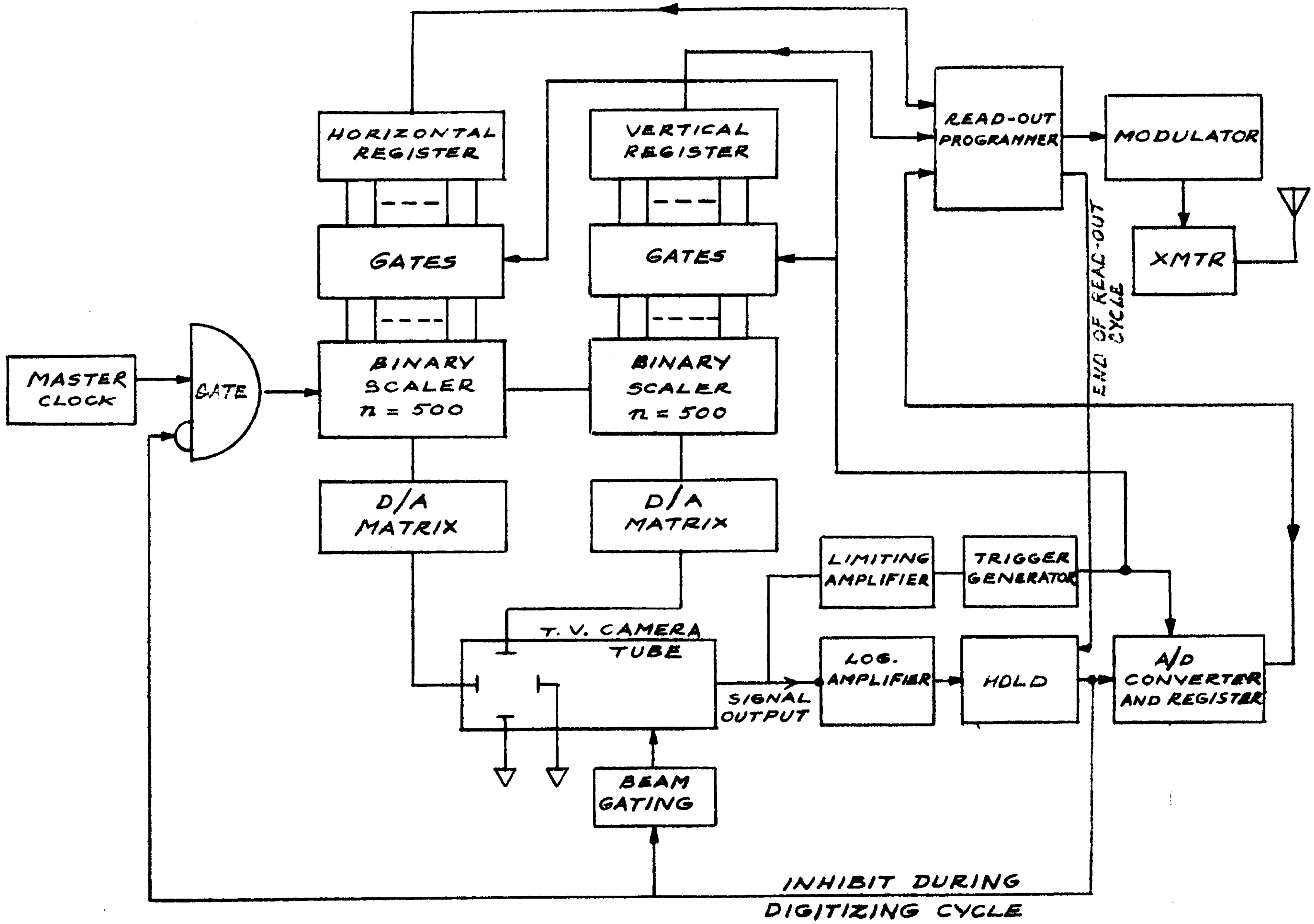


Figure XX. Digital TV System (Electrostatic Deflection) pulses are counted by the scaler, and the count is converted by a digital-to-analog (D/A) scaling matrix to a stair-step voltage which steps the beam across the face of the camera in increments corresponding to the resolution power of the tube. When the scaler reaches 500, it will reset to zero, causing the beam to flyback to its original position, and will put out a pulse to another scaler. The count of this scaler will be converted to a stair-step voltage in another D/A matrix, and this voltage will provide vertical deflection. In this manner the beam will scan the face of the tube, and the position of the beam at any time will be indicated by the count on the two scalers. Using the 250 kc master clock rate, it would take one second to scan an entire frame, assuming there were no output signals.

When the beam hits a star image, there will be a video output pulse, which will be amplified and fed to a hold circuit. The signal from the hold circuit will be used to close the gate from the master clock, thus stopping the scanning beam at this point. The signal from the hold circuit will also go to an analog-to-digital converter, thus providing a digital indication of the magnitude of the star. The amplifier must be a logarithmic amplifier, since magnitude is a logarithmic function of intensity. The video pulse will also be used to trigger gates to provide a non-destructive readout of the scalers into two registers. As soon as the registers are loaded, the read-out programmer will read them out in sequence and then reset the hold circuit, thus opening the gate from the master clock and starting the scan again. This process will be repeated each time a star is encountered. The signal that closes the gate should also turn off the scanning beam, lest the surrounding area of the mosaic be discharged by the beam resting on a single spot during the relatively long time required to read out the digital data.

Assuming straight binary coding for simplicity, the information per star is 23 bits -- 9 each for horizontal and vertical position and 5 for magnitude, so that the maximum information content of a frame is 3450 bits, and the average is about 1000 bits. With a 10 kc bit rate, this means a maximum frame time of about 1.35 sec and an average frame time of about 1.1 sec. If it is desired to cut this time down materially it will be necessary to increase the master clock frequency.

As mentioned earlier, there is also a requirement for television in the guidance of the satellite. The general procedure will be to use a small telescope with a 2 degree field as a part of closed loop positioning system which will operate to keep some particular star within a very small area of the field. There has been little or no specific design work done on this system, but a stop and go scanning system such as described above would seem to be suitable for this purpose.

The stop and go system described is certainly somewhat more complex than a conventional continuous-scan TV system, but the increase in complexity would not seem to be prohibitive in view of the large saving in bandwidth and required power. With the 10 kc bit rate specified, the bandwidth for return-to-zero coding would be 10 kc, which is a 15 to 1 advantage in bandwidth over conventional TV, for the same frame rate. Certainly a considerable amount of laboratory development would be needed to convert this basic system outline to a working system, but there is every reason to think the idea is practical. Indeed, the basic idea, in a somewhat simpler form, has already proven successful for facsimile transmission. Therefore it is recommended that further development of this system be made a part of the satellite program. It should be noted that several other systems of digital transmission of such star-field pictures have been considered, but the system described seems to be the only one that offers a good chance of successful development.

Another factor that should be considered briefly here is the ground station portion of the TV system. For the data-gathering TV there will be a requirement both for viewing and recording the images. Since this is continuous scan TV this requirement will involve a conventional receiver and a videotape recorder, with such modifications as may be necessary because of the slow scan. There is one special feature in connection with the recording: it will be necessary to place time reference signals on the tape. For this purpose the frame-sync pulses can be used to gate signals from the time signal generator. This is indicated in the block diagram of the ground station, Figure XIX.

Picture storage on the ground may be necessary with any slow speed system if the display is attempted with a conventional C.R.T. The persistence of a C.R.T., is in general not adequate to allow a picture to be presented at a slow speed, in that the portion of the picture first scanned begins to fade by the time the last portion is presented. This method of display would not allow an operator to view the field in its right perspective and would make identification extremely difficult, if not impossible. However, if the slow speed digital data is stored temporarily, then displayed at a higher repetition rate, C. R. T. persistence is not a problem.

The use of cathode ray storage tubes have also been considered for a display unit. Some of these storage tubes appear to have definite possibilities for an application such as this. The Hughes Tonotron, for example, has been used successfully as a display tube in slow scan closed circuit TV systems and as a radar PPI display scope. Because of the wide variations of brightness it presents, the Tonotron seems to be the most promising of the memory tube types. In most storage tubes the presentation has no brightness variation which would severely limit their usefulness in presenting star fields for identification. The Tonotron, however, sacrifices memory time for brightness variation. It has a resolution of 50 lines per inch which should be more than adequate and a storage time of approximately one minute. This storage time would allow several photographs to be taken of the tube presentation. Overlays, consisting of star fields on transparent backing could be used on the photographs to aid in identifying the star field.

For the stop-and-go finder TV, there will be a requirement for viewing the image. This will require a TV picture tube that will essentially reverse the sequence of operations described above in connection with the transmitter. There does not seem to be any need for recording the signals from this TV system, but if this were desired for some reason, a digital recorder such as the Ampex FR-400 would be required.

X. DATA STORAGE

The data storage requirements in the satellite are not severe when compared to the other technicological requirements of this program. However, the critical part played by a data storage device requires that careful consideration be given the overall problem. Special attention must be given to the reliability required of this device. Hence, tried and proven techniques are to be given preferential treatment.

Table III tabulates the five experiments being performed in the O.A.O. program and indicates the estimated maximum data storage requirements of each. From these data the maximum requirement is less than 10^6 bits during any one storage period, after which the storage device will presumably be read-out on command from the ground and reset, ready

TABLE III

Institution	Type of Experiment and Orbit	Experimental Equipment Requiring Data Storage	Maximum Bits Stored
U of Michigan	Solar spectrometry. Equatorial orbit	3 Spectrometers continuous scan 45 min. 20,000 info bits	140,000 (For 1% accuracy
Princeton	Stellar spectrometry. Polar orbit	Programmed scanning spectrograph	100,000 plus command program
Smithsonian	Sky mapping 3-color photometry Equatorial orbit	None	------------
NASA	Spectrophotometric measurements Equatorial orbit	Spectrograph	Experimenter will provide data storage
U of Wisconsin	Stellar spectrometry Equatorial orbit	Photomultipliers and ion detectors	Not large

to accept further data. If it is assumed that an acceptable time for clearing and resetting the register is 120 seconds, then 10^6 bits stored, a bit rate of approximately 10 kc would be adequate for transmission of the data. With these figures as tentative maxima, the merits of various storage devices have been examined and are discussed briefly below.

Magnetic Core Storage

The use of ferrite cores for storage of the data in the satellite has several distinct advantages; (1) they do not require moving parts for their operation (2) they can retain their memory throughout power failure (3) they are undoubtedly the most reliable electronic component in use in modern electronics.

Unfortuneately the disadvantages associated with using ferrite cores are as distinctive as their advantages where even moderate amounts of data storage are concerned. The read-in and read-out circuitry necessary to make use of core storage becomes extremely bulky and consumes a relatively large amount of power as the data volume increases beyond a few hundred bits. Further, when the speed of operation of the memory register becomes too fast for "mag-amps", as is the case at hand, then transistors and diodes must be used in the associated circuits in large numbers as the number of cores increase. Diodes in particular are not living up to their initial claims for reliability. The probability that a device, that requires a large number of diodes for its operation, will meet the longivity requirements of this satellite program, is not high.

Electrostatic Film Storage

RCA has recently developed a means for deposition charge patterns on an evacuated mylar film. The low lateral leakage attained with the mylar dielectric provides storage properties that are very desirable. Read-in and read-out equipment for this type storage device would not be large or power consuming. Further, relatively high speeds of operation could be achieved simply. Unfortunately, this storage device is still in the development stage and has not had sufficient use to justify choosing it for this program.

Magnetic Tape Recording

Magnetic tape recording is one of the oldest and most proven methods of data recording and storage in use today. However, the disadvantages that are immediately apparent from the standpoint of this program are the moving wheels necessary in the drive assembly and the possibility of tape oxide build-up on the recording heads when running unattended for long periods of time. The latter problem appeared to be the more serious of the two. This problem was discussed with several manufacturers of tape recorders, and with the manufacturers representatives for magnetic tape. In addition the experimental results were observed of a test program evaluating a magnetic tape recorder for the TIROS satellite. The conclusion drawn is that tape oxide build up with recorder running time is not a problem if the proper tape is used in a well designed tape transport. The problem is further eased if tape speeds are not excessive.

The moving parts in the recorder are a disadvantage that must be tolerated. However, careful attention to layout of the rotating parts would allow a large amount of the angular momentum to be counter-balanced so that the effect on the satellite stabilization system would not be large.

The power required to drive the recorder mechanism is dependent upon the speed of the recorder. The recorder would be run in a two speed cycle, low speed for recording and high speed for playback. The power required would be in direct proportion to the speed, however, the high speed cycle would be of short duration. Hence the total power drain would be tolerable.

If digital data is recorded on the tape, the signal-to-noise requirements are lessened with an attendant saving in the tape drive mechanism complexity. This is also compatible with the proposed digital telemetry system and would save data conversion equipment.

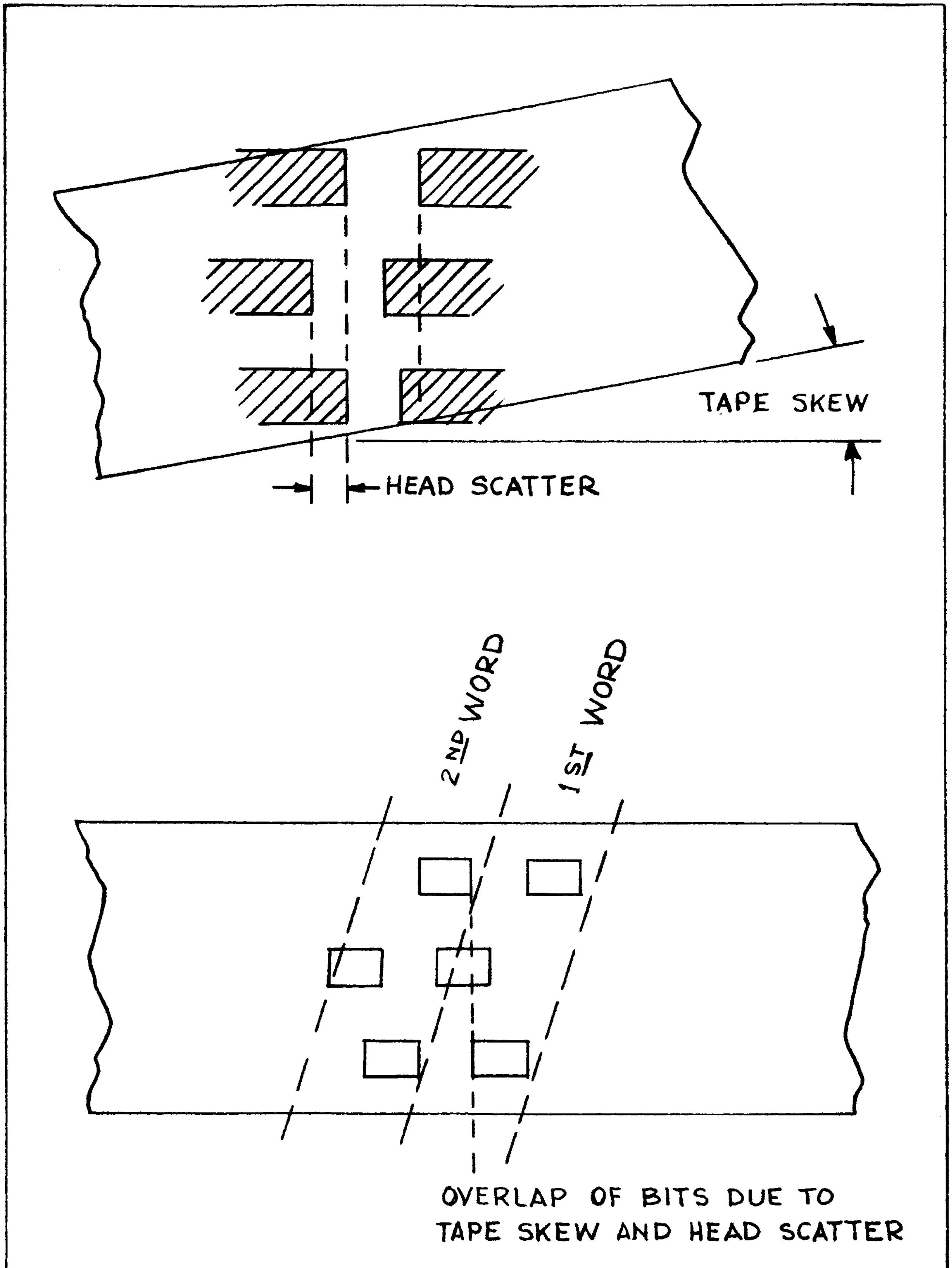
There are four generally accepted methods of recording digital data on magnetic tape. The RZ (return to zero) method uses (+) tape saturation as a "one" and (-) tape saturation as a "zero", returning to zero magnetic flux during the interval between a "one" and a "zero". RB (return to bias) method uses a bias current near negative tape saturation as an operating point. "Ones" are written by driving the tape to (+) saturation, "zeros" are not written. The RZ method is self-clocking in that a signal is available from each "one" and "zero" recorded. The RB must be gated to reestablish the code and therefore is not self-clocking. Either of the above methods may be used when high speed recording and maximum bit packing densities are not important. However, when it is necessary to put a greater number of bits per inch on the tape, one of the NRZ (non-return to zero) methods provide greatest econory. The NRZ-(M) (mark) method involves changing the state of the magnetization the tape from (+) saturation to (-) saturation or vice versa for each "one" in the code, "zeros" do not cause a change. The NRZ-(C) (change) method uses (+) saturation for "ones" and (-) saturation for "zeros"; however, during any sequence of "ones" or "zeros" the state of magnetization is not returned to a mid point. If three "zeros" are received in succession the state of magnetization remains at (-) saturation for the period of the three "zeros".

Both of the NRZ methods of recording require that the play-back signals be reshaped and gated to recover the coded information but have the advantage that at least twice the bit packing densities are allow-able over the RZ methods.

Head scatter and tape skew are two of the more important factors limiting the maximum allowable bit packing densities. These are in gen-eral determined by the mechanical design of the recorder being used. Figure XXI illustrates these defects. Head scatter is a result of the finite mechanical tolerances involved in keeping the recording heads for each track in exact alignment. The slight misalignment allows a scattering of the bits on the tape. If the scattering is not large compared to the distance between bits, the information can still be separated on playback. However, when the bits are packed too tightly, overlaps from one word to the next can occur and when this occurs the original information recorded cannot be resolved.

Tape skew, on the other hand results from the tape not being pulled in a uniform direction across the heads. If the angle between the direction of tape travel and the alignment of heads is not uniform, bit changes between recording and playback, overlap of the bits can occur.

These problems can be alleviated to a certain extent by re-cording one or more tracks of the tape clocking pulses that serve to



Effects of Tape Skew and Head Scatter
Figure XXI.

gate out the skewed relationship between pulses. This method is illustrated in Figure XXII.

It is desirable to record the clock pulses anyway for use when decoding the NRZ recording. In the manner illustrated above, two fold use is made of these recorded clock pulses. The above techniques allow bit packing densities of 500 bits/inch.

Imperfections in the magnetic tape are not as detrimental in digital recording as they are in analog recording, with one exception. Complete "drop-outs" where there is a hole in the oxide coating can cause loss or misinterpretation of a complete digital word. To alleviate this problem or at least to sense when this occurs, a parity check is usually made on each word and the parity bit recorded on one track of the tape. A simple parity check often used, is to count the number of "ones" in the digital word. If the number is even, a parity bit is generated and recorded. If the number is odd, no parity bit is recorded. The total number of "ones" across the tape for each word, including the parity bit, is always an odd number in this manner. A simple count of the number of "ones" in each recorded word serve to determine if the word sensed is the same as the word recorded. This scheme fails to sense correctly if an even number of "ones" are lost. If the importance of correct data warrants it, more complicated parity checks can be made to determine if a wrong word is sensed.

In a satellite magnetic tape data storage system a moderately complex digital record and playback system could be used that would record 400 bits/inch on the tape. With parallel word recording on a one inch tape 14 tracks are commonly used across the tape. If it is assumed that 10 of

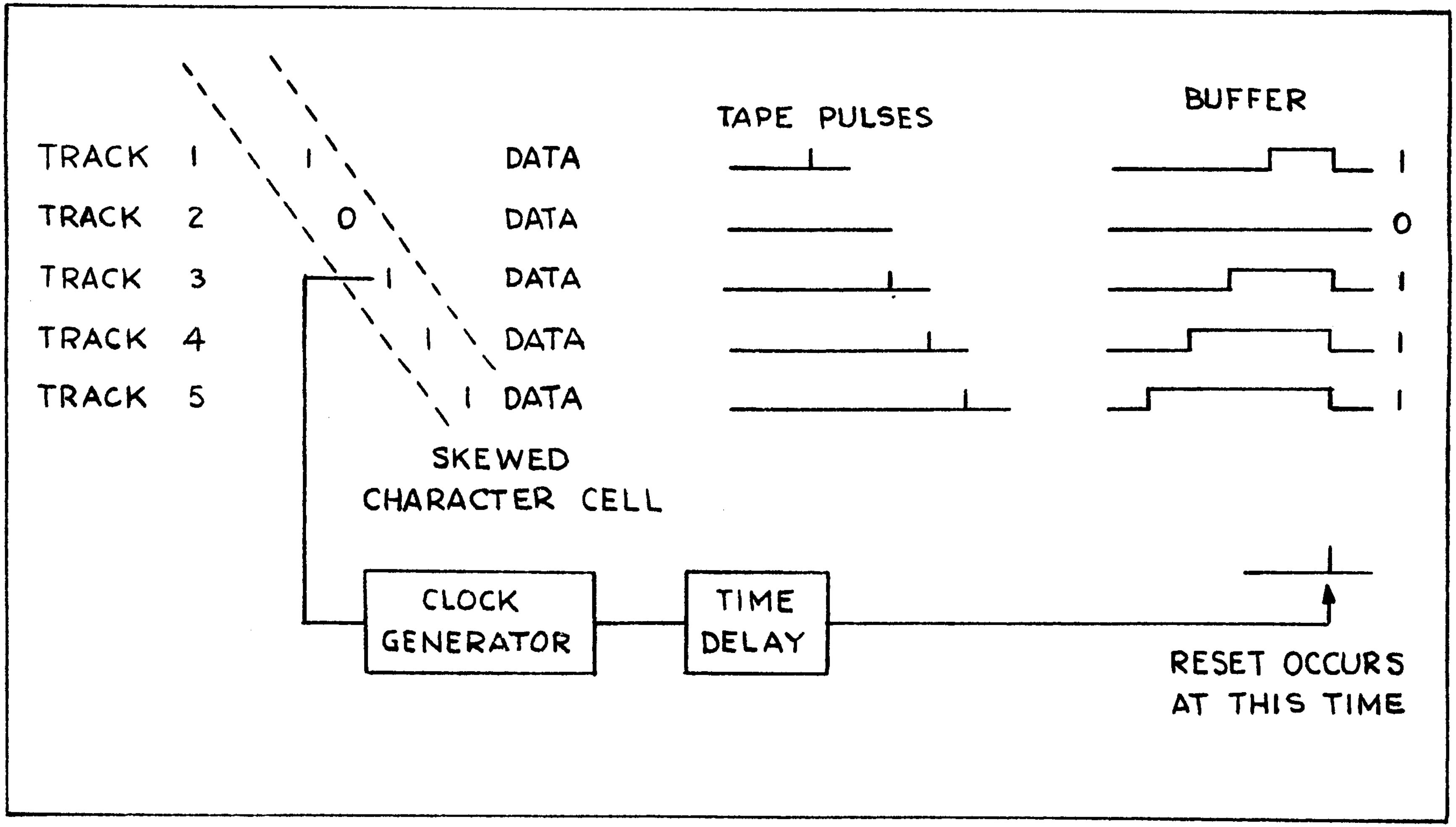


Figure XXII.
Recovery of Skewed Data

the tracks are used for data storage, the others being set aside for clocking pulses, etc., then a simple calculation shows that 10^6 bits would require only 250 inches (20.85 ft.) of tape.

It is obvious that the assumed digital recording figures above do not represent an optimum system. Less bit packing density would require longer tape but less complex circuitry to read-out the data. However, if the read-out time is held constant, then more tape must be pulled by the tape heads in this period. To accomplish this, higher tape transport speeds are necessary with higher resultant angular momentum and more power consumption. The final design will be a compromise between the conflicting requirements but will certainly be well within the state-of-the art.

XII BIBLIOGRAPHY

1. L. H. Aller, Astrophysics, the Atmosphere of the Sun and the Stars, Ronald Press, New York, 1953.

2. G. E. Anner, Elements of Television Systems, Prentice-Hall, New York, 1951.

3. G. F. Ault "Stable High Speed Digital to Analog Conversion for Storage Tube Deflection," Bell System Tech. J., Vol. 38, pp 445-465, 1959.

4. E. J. Baghdady, "F-M demodulator Time-Constant Requirements for Interference Rejection", Proc. I.R.E., pp 432-440, 1958.

5. J. E. Bartow, G. N. Krassner, and R. C. Riehs, "Design Considerations for Space Communication", 1959 I.R.E. Conv. Rec. pp 154-166. pt. 8.

6. W. H. Baum, "Counting Photons - One by One" Sky and Telescope, 14, P. 264, 1955.

7. F. K. Becker, J. R. Heffle and W. T. Wintringham, "An Experimental Visual Communication System," Bell System Tech. Journal, Vol 38, pp 141-176, 1959.

8. R. R. Bennett et al, "Circuits for Space Probes", Electronics, Vol. 32, pp 55-57, June 19, 1959.

9. H. S. Black, Modulation Theory, Van Nostrand, Princeton, N. J., 1953

10. K. Bullington, "Radio Propagation Fundamentals," Bell System Tech. Journal Vol. XXXVI, pp 593-644, 1957.

11. Caldwell, "Switching Circuits and Logical Design", Wiley, 1958 pp 405-411.

12. R. L. Choate, "Analysis of a Phase Modulation Communications System" ASTIA Document # AD234-529.

13. A. G. Clavier, P. F. Panter, and W. Dite, "Signal-to-Noise Ratio Improvement in a PCM System," Proc. I.R.E., Vol. 37, pp 355-359, 1949.

14. M. S. Corrington, "Variation of Bandwidth with Modulation Index in FM", Proc. I.R.E., Vol. 35, pp 1013-1020, 1947.

15. J. P. Costas, "Some Notes on Space Communications", Proc. I.R.E., Vol. 47, pp 1383, 1959.

16. J. P. Costas, "Synchronous Communications", Proc. I.R.E., Vol. 44, pp 1713-1718, 1956.

17. D. E. Cottrell, and R. E. Sturm, "A Research Study and Engineering Investigation of a Digital Television System", Quarterly Reports #1 and #2, ASTIA #AD206778 and #AD204326.

18. F. S. Coxe, "Operational Characteristics of the TLM-18 Automatic Tracking Telemetry Antenna", I.R.E. Trans., Vol. SET-5, pp 87-91, 1959.

19. M. G. Crosby, "Frequency Modulation Noise Characteristics", Proc. I.R.E., Vol. 25, pp 472-514, 1937

20. M. G. Crosby, "Frequency Modulation Propagation Characteristics," Proc. I.R.E., Vol. 24, pp 898-913, 1936.

21. M. G. Crosby, "The Service Range of Frequency Modulation," RCA Review, Vol. 4, pp 349-371, 1940.

22. H. B. Devore, "Limiting Resolution in an Image Orthicon Tube," Proc. I.R.E., Vol. 36, pp 335-345, 1948.

23. R. S. Djorup, "Comments Relative to the Application of PCM to Aircraft Flight Testing,". I.R.E. Nat. Conv. Rec. Pt. 5, pp 197-201, 1959.

24. Elias, "Predictive Coding", I.R.E. Transactions on Information Theory, Vol. IT-1, No.

25. H.-I. Ewen, "A Thermodynamic Analysis of Maser Systems," Microwave Journal Vol. 2, pp 41-46, March 1959.

26. J. S. Farrior, et al, "Attitude Control," Electrical Engineering, Col. 77 pp 1084-1090, 1958.

27. C. B. Feldman and W. R. Bennett, "Bandwidth and Transmission Performances," Bell System Tech. J., Vol. 28, pp 490-595, 1949.

28. P. Fellgett, "Theoretical and Practical Exploration of the Use of Television in Astronomy," Astronomical Optics, Z. Kopal, editor, Interscience, New York, 1956.

29. D. G. Fink, "Television Engineering, 2nd Ed., McGraw-Hill, New York, 1952.

30. D. G. Fink, editor, Television Engineering Handbook, McGraw-Hill, New York, 1957.

31. L. E. Flory, et al "Miniature ITV Camera Uses Drift Transistors," Electronics, Vol 30, pp 138- , Jan 1, 1958.

32. L. E. Flory, et al,"Television System for Stratoscope I," Electronics , Vol. 33, pp 49-53, June 17, 1960.

33. A. Gabor "Digital Signals at High Packing Densities", Electronics , October 16, 1959.

34. F. M. Gardner, "DOPLOC uses Phase-Locked Filter", Electronic Industries , Vol 18, No. 10, 1959.

35. Gilchrist, "The Application of Phase Locked Loop Discriminators for Threshold Improvement and Error Reduction in FM/FM Telemetry," 1957 National Telemetering Conference Report, pp IV-B-2.

36. W. M. Goodall, "Television by PCM, "Bell System Tech. J., pp 33-49, 1951.

37. G. G. Gouriet, "Bandwidth Compression of TV Signals," Proc. I.R.E., PtB, Vol. 104, pp 265-272, 1957.

38. Grabbe, Ramo and Wooldridge, Handbook of Automation, Computation and Control, Vol. I, Wiley, 1958. pp 16-12 and pp 36-38.

39. R. E. Graham, "Preditive Quantizing of Television Signals," I.R.E. Wescon Conv. Record, Pt. 4.

40. H. H. Grimm, "Fundamental Limitations of External Noise," I.R.E. Trans., Vol. I-8, pp 97-103, 1959.

41. Hamming, "Error Detecting and Correcting Codes", B.S.T.J. Vol. 29, pp 147-160, April 1950.

42. H. E. Haynes and D. T. Hoger, "Stop-and-Go Scanning Saves Spectrum Space," Electronics, Vol. 31, pp 84-88, Sept. 26, 1958.

43. D. A. Huffman, "Method for the Construction of Minimal - Redundancy Codes", Proceeding of the I.R.E. , Vol 40, 1098-1101, Sept. 1952.

44. R. Jaffe and E. Rechtin, "Design and Performance of Phase-Lock Circuits Capable of Near-Optimum Performance Over a Wide Range of Input Signal and Noise Levels," I.R.E. Trans., Vol. IT-1, pp 66-76, 1955.

45. Jet Propulsion Lab., "Transistorized Telemetering Development and Design Criteria," Report #20-108, 1957, ASTIA #AD130347.

46. H. Kihn and W. E. Barnette, "The Megacoder - A High Speed, Large-Capacity Microminiature Decoder for Selective Communication", RCA Review , Vol. 20 pp 153-179, 1959.

47. H. C. Ko, "The Distribution of Cosmic Radio Background Radiation," I.R.E. Proc., Vol. 46, pp 208-215, 1958.

48. E. R. Kretzmer, "Reduced Alphabet Representation of Television Signals," I.R.E. Nat. Conv. Rec., Pt. 4, pp 140-147, 1956.

49. V. D. Landon, "The Distribution of Amplitude with Time in Fluctuation Noise," I.R.E. Proc., Vol. 29 pp 50-55, 1941.

50. V. D. Landon and K. A. Norton, "Discussion on Distribution of Amplitude with Time in Fluctuation Noise," I.R.E. Proc., Vol 30 pp 425-429, 1942.

51. J. G. Lawton, "Comparison of Binary Data Transmission Systems," Proc., Nat. Conf. on Mil. Electronics, 1958.

52. I. M. Levitt, "The Motions of Satellites," Electrical Engineering, Vol. 77, pp 1078-1081, 1958.

53. D. Linden and A. F. Daniel, "New Power for Space-Age Electronics," Electronics, Vol. 32 pp 43-47, Mar. 20, 1959

54. B. D. Martin, "Threshold Improvement in an FM Subcarrier System", I.R.E. Trans Vol. SET-6, pp 25-33, 1960.

55. A. Maxwell, G. Swarup, and A. R. Thompson, "The Radio Spectrum of Solar Activity", I.R.E. Proc., Vol 46, pp 142-148, 1958.

56. C. T. McCoy, "Present and Future Capabilities of Microwave Crystal Receivers," I.R.E. Proc., Vol 46, pp 61-66, 1958.

57. J. D. McGee, "Photoelectric Aids in Astronomy," Astronomical Optics, Z. Kopal, editor, Interscience, New York, 1956.

58. D. D. McRae, "Consideration of RF Parameters for PCM Telemetry Systems," I.R.E. Trans., Vol. SET-5, pp 61-65, 1959.

59. D. D. McRae, "Phase-Locked Demodulation in Telemetry Receivers," I.R.E. Proc. NTL Symposium on Telemetering, 1958.

60. G. F. Montgomery, "A Comparison of Amplitude and Angle Modulation for Narrow-Band Communication of Binary-Coded Messages in Fluctuation Noise," I.R.E. Proc. Vol 42, pp 447-454, 1954.

61. M. H. Nichols and R. L. Rauch, Radio Telemetry, Wiley, New York 1956.

62. T. R. Nisbet, "How Radiation Affects Electronic Equipment," Electronic Equipment Engineering, Vol. 7, pp 36-40 May, 1959.

63. A. P. Norton, Star Atlas, Gall and Inglis, London, 1946.

64. B. M. Oliver, "Efficient Coding," Bell System Tech. J. Vol. 31, pp 724-750, 1952.

65. B. M. Oliver, J. R. Pierce, and C. E. Shannon, "The Phiosophy of PCM," I.R.E. Proc., Vol. 36, pp 1324-1331, 1948.

66. "Radio Propagation Factors Which Influence the Selection of Frequencies for Outer Space Communications," ASTIA #AD201378.

67. V. A. Ratner, "Telemetry Diversity Reception Techniques," I.R.E. Proc., Nat. Symposium of Space Electronics and Telemetry, 1959.

68. S. O. Rice, "Mathematical Analysis of Random Noise," Noise and Stochastic Processes,

69. H. L. Richter, W. F. Sampson, R. Stevens, "Microlock: A Minimum Weight Radio Instrumentation System for a Satellite," Proc. 1957 Nat. Telemetering Conference.

70. R. W. Rochelle, "Earth Satellite Telemetry Coding System," Electrical Engineering, Vol. 76, pp 1062-1065, 1957.

71. W. F. Sampson and F. A. Ruegg, "Phase-Lock in Space Communications," I.R.E. Proc., Nat. Symposium of Space Electronics and Telemetry, 1959.

72. R. W. Sanders, "Digilock Telemetry System," I.R.E. Proc., Nat. Symp. on Space Electronics and Telemetry, 1959.

73. O. H. Schade, "Electro-Optical Characteristics of Television Systems," RCA Review, Vol. 9, pp 5, 245, 490, 653, 1948.

74. D. H. Schaefer, "Magnetic Core Event Counter for Earth Satellite Memory," Electrical Engineering, Vol. 77, pp 52-56, 1958.

75. W. F. Schreiber and C. F. Knapp, "TV Bandwidth Reduction by Digital Coding," I.R.E. Nat. Conv. Rec., Pt. 4, pp 88-99, 1958.

76. Mischa Schwartz, "Information Transmission, Modulation, and Noise", McGraw-Hill, 1959.

77. F. H. Seares and M. C. Joyner, "Systematic Deviations from the Mean Stellar Distribution," Contr. Mt. Wilson Obs., No. 346, 1928.

78. W. H. Seaver, "PCM Synchronization", Proc. Ntl. Telemetering Conference, 1960.

79. C. E. Shannon, "Communication in the Prescence of Noise," Proc. I.R.E., Vol. 37, pp 10-21, 1949.

80. K. R. Spangenberg, Vacuum Tubes, McGraw-Hill, New York, 1948.

81. "Standards for PCM Telemetry," I.R.E. Trans., Vol. SET-5, pp 194-195, 1959.

82. H. P. Steier, "Spaceship Telemetry," Missles and Rockets, Vol. 3, pp 105-110, 1958.

83. O. Struve, "Some Astronomical Applications of Image Converters", Sky and Telescope, 14, pp 224, 1955.

84. E. D. Sunde, "Theoretical Fundamentals of Pulse Transmission" Bell System Tech. J., Vol. 33, pp 721-788, 987-1010, 1954.

85. D. E. Sukhia, "Space Vehicle Communications System Design," Proc., Nat. Conv. on Mil. Electronics, 1958

86. "Telemetry Frequency Utilization Parameters and Criteria," IRIG Recommendation #101-59, Jan., 1959.

87. Telemetry Standards, "IRIG Document #103-56, Oct.,1956.

88. B. Thompson, "Data Processing with Magnetic Tape Recorders," I.R.E. Proc., Nat. Symposium on Telemetering, 1958.

89. R. J. Trumpler and D. F. Weaver, Statistical Astronomy, Univ. of Calif. Press, Berkeley, 1953.

90. C. S. Weaver, "A New Approach to the Linear Design and Analysis of Phase-Locked Loops," IRE Trans., Vol. SET-5 pp 166-78, 1959.

91. W. L. Wilcox and J. E. Geake, "The Design and Operation of a Photoelectric Spectro-Photometer," Astronomical Optics, Z. Kopal, editor, Interscience, New York, 1956.

92. M. W. Williard and V. A. Ratner, "A Universal PCM Data Collection System," Proc. Ntl. Telemetering Conference, 1960.

93. L. P. Yeh, "Communicating in Space," Electronic Industries, Vol. 18, pp. 54-58, Feb. 1959 and pp 94-97, Mar. 1959.

94. V. K. Zworykin and G. A. Morton, Television, 2nd Ed., Wiley, New York, 1954.